

BANDIT PROBLEMS WITH LÉVY PAYOFF PROCESSES

ASAF COHEN AND EILON SOLAN

ABSTRACT. We study two-armed Lévy bandits in continuous-time, which have one safe arm that yields a constant payoff s , and one risky arm that can be either of type High or Low; both types yield stochastic payoffs generated by a Lévy process. The expectation of the Lévy process when the arm is High is greater than s , and lower than s if the arm is Low.

The decision maker (DM) has to choose, at any given time t , the fraction of resource to be allocated to each arm over the time interval $[t, t + dt)$. We show that under proper conditions on the Lévy processes, there is a unique optimal strategy, which is a cut-off strategy, and we provide an explicit formula for the cut-off and the optimal payoff, as a function of the data of the problem. We also examine the case where the DM has incorrect prior over the type of the risky arm, and we calculate the expected payoff gained by a DM who plays the optimal strategy that corresponds to the incorrect prior.

In addition, we study two applications of the results: (a) we show how to price information in two-armed Lévy bandit problem, and (b) we investigate who fares better in two-armed bandit problems: an optimist who assigns to High a probability higher than the true probability, or a pessimist who assigns to High a probability lower than the true probability.

1. INTRODUCTION

Consider a firm that has to determine, on an ongoing basis, how much to invest in the research of new technologies for its next line of products. The firm faces a tradeoff between exploration and exploitation: on the one hand, it can adopt the technology that seems most successful according to the research conducted so far, thereby exploiting its investment in research, but on the other hand, it could continue investing in various technologies, in the hope of finding an even better technology for its products. If the firm decides to stop investing in a given technology, then no information will be obtained on that technology, so even if it is actually better than the finally adopted technology, it will never be adopted.

A similar tradeoff between exploration and exploitation arises, e.g., in the market of venture capital funds, where each fund has to decide in which start-up companies to invest, and in clinical trials, where pharmaceutical companies have to decide which new drugs or treatments to explore.

To concentrate on the trade off between exploration and exploitation, one assumes that there are no exogenous factors that affect the firm's decision (such as

Key words and phrases. Two-armed Bandit, Levy processes, cut-off strategies.

The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: cohenasa@post.tau.ac.il.

The School of Mathematical Sciences, Tel Aviv University, Tel Aviv 69978, Israel. e-mail: eilons@post.tau.ac.il.

We thank Adi Ditkowski and Steven Schochet for their helpful remarks about differential equations, and Sven Rady for useful comments on an earlier version of the paper.

new technologies or drugs that are introduced by competitors). The optimization problem that the firm faces has been modeled in the literature as a multi-arm bandit problem (see, e.g., Rothschild (1974), Bergemann and Välimäki (2006), Keller, Rady and Cripps (2005), Besanko and Wu (2008), Klein and Rady (2008), Moscarini and Squintani (2004), Roberts and Weitzman (1981), Weitzman (1979)): a decision maker (DM) has finitely many actions, called arms, each one yields a payoff with an unknown distribution, that is taken from a finite set of distributions. Each time the DM chooses an arm, he obtains a payoff, and improves his information regarding the correct payoff distribution of the arm he has just chosen.

Gittins and Jones (1979) proved that, in discrete-time, the optimal strategy of the DM has a particularly simple form: at every period the DM calculates a real number, an index, for each arm, based on past observations of that arm, and he chooses the arm with the highest index. It turns out that to calculate the index of an arm it is sufficient to consider an auxiliary problem with two arms: the arm for which we calculate the index, and an arm that yields a constant payoff. The literature therefore focuses on such problems, called two-armed bandit problems.

Once the optimality of the index strategy is guaranteed, one looks for the relation between that data of the game and the index. Explicit formulas for the index when the payoff is one of two distributions that have a simple form has been established in the literature. Berry and Fristedt (1985) provide the solution to the problem in discrete-time, e.g., when the payoff distribution is one of two Bernoulli distributions, and in continuous-time, e.g., when the payoff distribution is one of two Brownian motions. In continuous-time, by studying the dynamic programming equation that describes the problem, Keller, Rady and Cripps (2005) and Keller and Rady (2008) provided an explicit form for the index when the payoff's distribution is Poisson.¹ When the payoff distribution is known, Karatzas (1984) characterized the index when the payoff's distribution is a diffusion process, and Kaspi and Mandelbaum (1995) characterized the index when the payoff's distribution is a Lévy process, and they obtained an explicit form for the index for special distributions.

In practice, payoff processes have a complex form, exhibiting both small random changes, that can be modeled by a Brownian motion, and large shocks that can be modeled as arriving at a Poisson rate. A stochastic process that incorporates these two types of changes is the Lévy process. In fact, Carr and Wu (2004) argue that almost all economic phenomena can be described by time shifts of Lévy processes. Therefore, it is desirable to study the bandit problem where the payoff distribution is one of finitely many Lévy distributions.

In the present paper we provide an explicit solution to the two-armed bandit problem where the payoff distribution is one of two Lévy processes. We assume that one distribution, called *High*, dominates the other, called *Low*, in a strong sense (see Assumption 2.1 below). To eliminate trivial cases, we assume that the expected payoff that is generated by the safe arm is lower than the expected payoff generated by the High distribution, and higher than the expected payoff generated by the Low distribution.

In such a case in discrete-time, the optimal strategy is a cut-off strategy: the DM keeps on experimenting as long as the posterior belief that the distribution is

¹These authors also studied the strategic setup, in which several DMs have the same set of arms, and their arms' payoff distributions are the same (and unknown), and they compared the cooperative solution to the non-cooperative solution.

High is higher than some cut-off point, and, once the posterior probability that the distribution is High falls below the cut-off point, the DM switches to the safe arm. We prove that when the two payoff distributions are Lévy processes that satisfy several requirements, the optimal strategy is a cut-off strategy, and we provide an explicit expression for the cut-off point, in terms of the data of the problem. When particularized to the models studied by Kaspi and Mandelbaum (1995), Bolton and Harris (1999), Keller, Rady and Cripps (2005) and Keller and Rady (2008), our expression reduces to the expressions that they obtained.

Apart of unifying previous results, our characterization shows that the special form of the optimal payoff derived by Bolton and Harris (1999) and Keller, Rady and Cripps (2005) is valid in a general setup: the optimal payoff is the sum of the expected payoff if no information is available, and of an option value, that measures the expected gain from the ability to experiment. It also shows that the data of the problem can be divided into information-relevant parameters and payoff-relevant parameters; the information-relevant parameters can be summarized in a single real number, and the payoff-relevant parameters are the expectations of the processes that contribute to the DM's payoff. Finally, the characterization allows one to derive comparative statics on the optimal cut-off and payoff. For example, as the discount rate increases, or the signals become less informative, the cut-off point increases but the DM's optimal payoff decreases.

It is often the case that the DM holds an incorrect prior, and plays optimally given that prior (for further discussion and literature review, see Section 2.4). We provide an explicit expression for the value function in this case. Our technique provides a new description of the optimal strategy with a time dependent cut-off.

So far we have assumed that all the information that the DM has is the payoff process. Sometimes, the DM has information that does not contribute to the payoff process, yet it helps him learn the type of risky arm. For example, scientific discoveries made by other firms in other markets may shed light on the appropriateness of a given technology to a product that a firm develops. In addition, part of the payoff of the DM may not be observed by the DM.

In Section 2.5 we study a bandit problem, in which the risky arm generates three Lévy processes - the first is observed by the DM and contributes to the payoff, the second is observed by the DM and does not contribute to the payoff, and the third is not observed by the DM but it contributes to the payoff. We provide an explicit expression for the cut-off and for the optimal payoff of the DM. That generalizes the expression we found when the risky arm generates only one process. This analysis clarifies the distinction between information-relevant data and payoff-relevant data.

We conclude the paper by applying our characterization to compare the effects of optimism and pessimism in bandit problems. A DM is called optimist if his prior probability that the payoff's distribution is High is higher than the true probability, and he is called a pessimist if his prior probability that the payoff's distribution is High is lower than the true probability. Using our characterization we find that unless the pessimist assigns high probability to the High type and the two DM's are sufficiently patient, an optimist will fare better than a pessimist.

The rest of the paper is arranged as follows. The model and the main results appear in Section 2. Directions for future research appear in Section 3. All proofs appear in Section 4.

2. THE MODEL AND THE MAIN RESULTS

2.1. Reminder about Lévy Processes. Lévy processes are the continuous-time analog of discrete-time random walks with i.i.d. increments. A *Lévy process* $X = (X(t))_{t \geq 0}$ is a continuous-time stochastic process that (a) starts at the origin: $X(0) = 0$, (b) admits càdlàg modification,² and (c) has stationary independent increments. See figures 1 and 2 for a generic path of a Lévy process. A few examples of Lévy processes are a Brownian motion, a Poisson process, and a compound Poisson process. The latter is a continuous-time process in which jumps arrive according to a Poisson process and the jumps are i.i.d.³

We now present the Lévy-Ito decomposition of Lévy processes. Let $(X(t))$ be a Lévy process. For every Borel measurable set $A \subseteq \mathbb{R} \setminus \{0\}$, define:

$$\nu(A) := E[\#\{0 \leq s \leq 1 | \Delta X(s) := X(s) - X(s-) \in A\}].$$

This is the expected number of jumps with size in A that occurs up to time 1. One can verify that ν is a measure, called the *Lévy measure* of the process. The quantity $\nu(\mathbb{R} \setminus \{0\})$ is the expected number of jumps that occur up to time 1. If $\nu(\mathbb{R} \setminus \{0\}) = \infty$, then the number of jumps in the time interval $[0, 1]$, and therefore in any compact time interval, is infinite a.s., and we say that the Lévy measure is infinite. If $\nu(\mathbb{R} \setminus \{0\}) < \infty$, then the expected number of jumps in any compact time interval is finite a.s., and we say that the Lévy measure is finite. In this paper we study Lévy processes that have finite Lévy measures. The results can be generalized for Lévy processes with infinite Lévy measures, see Cohen and Solan (2009).

The Lévy-Ito decomposition (see Applebaum (2004)) states that every Lévy process with finite Lévy measure can be represented as follows:

$$(2.1) \quad X(t) = \mu t + \sigma Z(t) + L_\nu(t),$$

where μt is a linear drift, $\sigma Z(t)$ is a Brownian motion with standard deviation σ , and $L_\nu(t)$ is a compound Poisson process with Lévy measure ν which is independent of $\sigma Z(t)$: jumps arrive at a Poisson rate with expectation $\nu(\mathbb{R} \setminus \{0\})$, and the distribution of each jump is given by the distribution function $\frac{\nu(dh)}{\nu(\mathbb{R} \setminus \{0\})}$.

2.2. Lévy Bandits. A DM operates a two-armed bandit machine in continuous-time, with a safe arm that yields a constant payoff s , and a risky arm that yields a stochastic payoff $(X(t))$. The risky arm can be of two types, High or Low. We denote the arm's type by θ : if the type is High (resp. Low) we set $\theta = \theta_1$ (resp. θ_2). If $\theta = \theta_i$, $i \in \{1, 2\}$, the risky arm yields payoff $(X_i(t))$, which is a Lévy process. We assume throughout that the Lévy measures of both $(X_1(t))$ and $(X_2(t))$ are finite, and therefore a.s. there are only finitely many jumps in each compact time interval. Denote the Lévy-Ito decomposition of $(X_i(t))$, $i \in \{1, 2\}$, by $X_i(t) = \mu_i t + \sigma_i Z(t) + L_{\nu_i}(t)$.

²That is, it is continuous from the right, and has limits from the left: for every t_0 , the limit $X(t_0-) := \lim_{t \nearrow t_0} X(t)$ exists a.s. and $X(t_0) = \lim_{t \searrow t_0} X(t)$.

³Formally, let $\lambda > 0$, and let D be a distribution over $\mathbb{R} \setminus \{0\}$. A *compound Poisson process* with rate λ and jump size distribution D is a continuous-time stochastic process given by $X(t) = \sum_{i=1}^{N(t)} D_i$, where $N(t)$ is a Poisson process with rate λ , and D_i are i.i.d. random variables, with distribution function D , which are also independent of $(N(t))_{t \geq 0}$.

Set $\bar{\nu}_i := \nu_i(\mathbb{R} \setminus \{0\})$, and denote by $H_i := \int h\nu_i(dh)/\bar{\nu}_i$ the expected jump size of $(X_i(t))$.⁴ We assume that H_i is finite. The quantity $\int h\nu_i(dh) = \bar{\nu}_i H_i$ is the contribution of the compound Poisson process to the instantaneous payoff. The expectation of the risky arm at time $t = 1$ if $\theta = \theta_i$ is $g_i := E[X_i(1)] = \bar{\nu}_i H_i + \mu_i$.

Throughout we make the following assumption, which states that the High type is better than the Low type in a strong sense.

Assumption 2.1.

A1. $-\infty < g_2 < s < g_1 < \infty$.

A2. $\sigma_1 = \sigma_2$.

A3. For every $A \in \mathcal{B}(\mathbb{R} \setminus \{0\})$, $\nu_2(A) \leq \nu_1(A) < \infty$.

Assumptions A1 and A2 rule out trivial cases. Assumption A1 merely says that the High (resp. Low) type provides higher (resp. lower) expected payoff than the safe arm. Assumption A2 states that the Brownian motion component of both the High type and Low type have the same standard deviation. Otherwise, since the realized path reveals the standard deviation, the DM can distinguish between the arms in any infinitesimal time interval.

The third part of the assumption is less innocuous; it requires that the Lévy measure of the High type will dominate the Lévy measure of the Low type in a strong sense: roughly, jumps of any size h occur more often (or at the same rate) under the high type than under the Low type.

A consequence of this assumption is that jumps always provide good news, and (weakly) increase the posterior probability of the High type.

At each time instance t , the DM chooses the proportion of time to devote to each of the two arms. If he chooses to devote a proportion k of the current time instance to the risky arm, then the instantaneous payoff dY^k is the sum of several terms:

- $(1 - k)sdt$, which is the contribution of the safe arm;
- $k\mu_i dt$, which is the contribution of the linear drift;
- $\sqrt{k}\sigma dZ(t)$, which is the contribution of the Brownian motion;⁵
- $kH_i dt$, which is the contribution of the compound Poisson process.

A strategy κ is a (measurable) function, that assigns to each history a number in the interval $[0, 1]$, that is interpreted as the amount of time in the interval $[t, t + dt)$ devoted to the risky arm. In continuous time, it is usually assumed that a strategy is predictable, that is, to determine the behavior at time t , it is sufficient to know the history strictly before time t . Formally, κ_t is \mathcal{F}_{t-}^I -measurable, where $I(t) := \int_0^t re^{-rt} dY^\kappa(t)$ is the stochastic discounted payoff using the strategy κ , and \mathcal{F}_{t-}^I is the σ -algebra generated by the stochastic process $(I(t))$ of the discounted payoff with discount rate r , up to (excluding) time t .

As is well known, in continuous-time the play path need not be uniquely defined, and therefore one should be careful in defining the set of strategies available to the DM, and the play path that a strategy defines. We circumvent this issue by arguing that an optimal strategy must solve a certain Functional Differential Equation

⁴In order to simplify notation, we denote $\int h\nu_i(dh) := \int_{\mathbb{R} \setminus \{0\}} h\nu_i(dh)$.

⁵By devoting the proportion κ of the time interval $[t, t + dt)$ to the risky arm, the variance of the continuous part of the payoff is $\kappa\sigma^2 dt$. This explains the scaling parameter $\sqrt{k}\sigma$. For a qualitative explanation for this form, see Bolton and Harris (1999).

(FDE), by showing that this FDE has a unique solution, and by exhibiting this solution. As for discrete-time, this solution turns out to be a cut-off strategy, which we now define.

Let $p_t := P(\theta = \theta_1 | \mathcal{F}_{t-}^I)$ be the posterior belief at time t that the risky arm is High. A strategy κ is a *cut-off strategy* with cut-off point p if the DM chooses the safe arm (with probability 1) whenever $p_t \leq p$, and the risky arm (with probability 1) whenever $p_t > p$. We now argue that Assumption A3 guarantees that the play under a cut-off strategy is well defined. Suppose that $p_0 > p$. Then the DM chooses the risky arm until the first time t that satisfies $p_t = p$. Since the Lévy payoff processes have finite measures, Assumption A3 implies that in an infinitesimal interval after time t the posterior belief will drop below the cut-off. Indeed, the first time t satisfying $p_t = p$ is a predictable stopping time, and therefore, the probability that a jump occurs at that time is zero (see Bertoin (1996)). Therefore, $P(\theta = \theta_1 | \mathcal{F}_t^I) = P(\theta = \theta_1 | \mathcal{F}_{t-}^I) = p$. If there is a Brownian motion component, its fluctuations will cause the posterior to drop below p in an infinitesimal time interval after time t . If there is no Brownian motion component, since the Lévy measure is finite, there are no jumps in an infinitesimal time interval, and by Assumption A3 the compound Poisson process will cause the posterior to drop below p . This implies that the play under a cut-off strategy is well defined.

2.3. The Optimal Strategy. The expected discounted payoff under a strategy κ when the prior is $p_0 = p$ is

$$\begin{aligned} V_\kappa(p) &= E \left[\int_0^\infty r e^{-rt} dY^\kappa(t) \right] \\ &= p E \left[\int_0^\infty r e^{-rt} dY^\kappa(t) \middle| \theta_1 \right] + (1-p) E \left[\int_0^\infty r e^{-rt} dY^\kappa(t) \middle| \theta_2 \right]. \end{aligned}$$

Let $U(p) = \sup_{\kappa} V_\kappa(p)$ be the maximal payoff the DM can achieve. As we show below, the DM has an optimal strategy, so in fact the supremum in the definition of $U(p)$ is achieved. The function $V_\kappa(p)$ is linear with respect to p , and therefore $U(p)$, as the supremum of linear functions, is convex. By always choosing the safe arm, the DM can achieve at least s ; since $U(0) = s$, the convexity of $U(p)$ implies that U is non-decreasing.

Proposition 2.2. *$U(p)$ is monotone non-decreasing, convex, and continuous in p .*

It follows from Proposition 2.2 that there is p^* such that $U(p) = s$ if $p \leq p^*$ and $U(p) > s$ otherwise, so that the strategy $\kappa \equiv 0$ that always chooses the safe arm is optimal for prior beliefs in $[0, p^*]$.

Our first theorem states that there is a unique optimal strategy, which is a cut-off strategy. Moreover it provides the exact cut-off point and the corresponding expected payoff in terms of the data of the problem. Let α be the unique solution of

$$(2.2) \quad f(\eta) := \int \nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\bar{\nu}_1 - \bar{\nu}_2) - \bar{\nu}_2 + \frac{1}{2}(\eta+1)\eta \left(\frac{\mu_1 - \mu_2}{\sigma} \right)^2 - r = 0$$

in $(0, \infty)$. The existence and uniqueness of such a solution are proved in Lemma 4.4 below. Observe that $\frac{\nu_2(dh)}{\nu_1(dh)}$, the Radon-Nikodym derivative, exists by Assumption A3.⁶

Theorem 2.3. Denote $p^* := \frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$. Under Assumptions A1-A3, the unique optimal strategy is

$$\kappa^* = \begin{cases} 0 & \text{if } p \leq p^*, \\ 1 & \text{if } p > p^*. \end{cases}$$

The expected payoff under κ^* is

$$(2.3) \quad U(p) = V_{\kappa^*}(p) = \begin{cases} s & \text{if } p \leq p^*, \\ pg_1 + (1-p)g_2 + C_\alpha(1-p)\left(\frac{1-p}{p}\right)^\alpha & \text{if } p > p^*, \end{cases}$$

$$\text{where } C_\alpha = \frac{s-g_2-p^*(g_1-g_2)}{(1-p^*)\left(\frac{1-p^*}{p^*}\right)^\alpha}.$$

The term $pg_1 + (1-p)g_2$ in (2.3) is the expected payoff for a DM choosing only the risky arm. Thus, $C_\alpha(1-p)\left(\frac{1-p}{p}\right)^\alpha$ is the option value for the ability to switch to the safe arm. The quantity α summarizes all the information-relevant parameters of the problem (see also Section 2.5). Apart from that, the parameters of the payoff processes that determine the cut-off point p^* and the optimal payoff $U(p)$ are the expected payoffs g_1 and g_2 .

The function in (2.3) has the same structure as the solution of Bolton and Harris (1999) and Keller, Rady and Cripps (2005) for the one agent problem. In Bolton and Harris (1999), the only component in the risky arm is the Brownian motion with drift. Therefore $\nu_i \equiv 0$, so that $g_1 = \mu_1$, $g_2 = \mu_2$, and $\alpha = (-1 + \sqrt{1 + 8r\sigma^2/(\mu_1 - \mu_2)^2})/2$. In Keller, Rady and Cripps (2005), the risky arm is either the constant zero (Low type, so that $\nu_2 \equiv 0$), or yields a payoff \bar{h} according to a Poisson process with rate λ (High type). If the risky arm is High, the only component in the Lévy-Ito decomposition is the compound Poisson component, and $\nu_1(\bar{h}) = \lambda$ and zero otherwise. Therefore, $g_1 = \lambda\bar{h}$, $g_2 = 0$, and $\alpha = r/\lambda$.

The explicit form of the cut-off point p^* and of the value function U allows us to derive simple comparative statics. As is well known, a DM who plays optimally switches to the safe arm later than a myopic DM, and indeed, p^* is smaller than the myopic cut-off point $p^m := \frac{s-g_2}{g_1-g_2}$. Furthermore, the cut-off point p^* is an increasing function of α . As can be expected, α (and therefore also p^*) increases at the discount rate r and at $\nu_2(dh)$, and it decreases at $\nu_1(dh)$ and at $|\mu_1 - \mu_2|$: the DM switches to the safe arm earlier as the discount rate increases, as jumps provide less information, or as the difference between the drifts of the two types increases.⁷ Furthermore, as long as $p > p^*$, the value function $U(p)$ decreases in α . Thus, decreasing the discount rate, increasing the informativeness of the jumps and the difference between the drifts is beneficial to the DM.

⁶To ensure the existence of the Radon-Nikodym derivative, and therefore the form of the solution, one does not need the full power of Assumption A3. Its full power will be used within the proof.

⁷Moreover, $\alpha(r=0) = 0$ and $\alpha(r=\infty) = \infty$.

In Section 2.4 we extend the results to the case that the prior belief of the DM is not the true prior p_0 . In Sections 2.6 and 2.5 we provide two applications to our techniques and results.

Remark 2.4. Incomplete information of p_0 . Suppose that the DM does not know the prior belief p_0 , but rather has some belief φ over p_0 . That is, p_0 is chosen at the outset according to φ , and is not told to the DM. From the DM's point of view, the situation is equivalent to an auxiliary problem in which the prior probability of the High type is the expectation of the corresponding probability in the original problem, and therefore Theorem 2.3 provides the optimal strategy in this case as well.

2.4. The Payoff with Incorrect Prior. In decision problems it is usually assumed that the DM either knows the true state of nature, or has some prior distribution over the set of states of nature. Experiments show that the prior distribution that DMs have is often different than the true prior. The phenomenon of overconfidence – assigning too high a probability to the good state of nature – has been observed in various areas (Svenson (1981), Baumhart (1968), Larwood and Whitaker (1977), Cross (1977), Weinstein (1980), Camerer and Lovallo (1999)). Babcock and Loewenstein (1997) argue that biases in bargaining may be self-serving, and Heifetz, Shannon and Spiegel (2007) show that biases of preferences may be stable.

In every decision problem, a DM who correctly perceives the prior distribution will fare better than a DM who has some bias, and believes that the prior distribution is different than the correct one. Indeed, an optimal strategy of a DM who correctly perceives the prior distribution is a strategy that yields the highest possible gains for this prior distribution, so it yields at least as much as any other strategy, in particular, optimal strategies for incorrect prior distributions.

Denote the initial belief of the DM for the High type by q_0 , and suppose that it may be different from the true probability p_0 . By Theorem 2.3, the optimal strategy of the DM is a cut-off strategy, however, since he has an incorrect prior, he does not switch to the safe arm at the optimal time. In this section we give an exact formula for the payoff, assuming the DM plays optimally given his belief. We will also describe the optimal strategy from a different point of view, not as a cut-off strategy. This point of view is arguably closer to the way people perceive the decision problem that the DM faces.

Suppose that until time t , the DM chose the risky arm, and observed the jumps h_1, \dots, h_n from the compound Poisson component of the payoff process. Let $Y_B^1(t)$ be the Brownian motion with drift component of the payoff process from the risky arm at time t . Note that $Y_B^1(t) \sim N(\mu t, \sigma^2 t)$. The posterior belief

$q_t := P_t(\theta_1 | h_1, \dots, h_n; Y_B^1(t); q_0)$ of the DM is:⁸

(2.4)

$$\begin{aligned} q_t &= \frac{q_0 \frac{1}{\sqrt{2\pi t \sigma}} e^{-\frac{(Y_B^1(t) - \mu_1 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_1 t} \prod_{t-} \nu_1(dh_j)}{q_0 \frac{1}{\sqrt{2\pi t \sigma}} e^{-\frac{(Y_B^1(t) - \mu_1 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_1 t} \prod_{t-} \nu_1(dh_j) + (1 - q_0) \frac{1}{\sqrt{2\pi t \sigma}} e^{-\frac{(Y_B^1(t) - \mu_2 t)^2}{2\sigma^2 t}} e^{-\bar{\nu}_2 t} \prod_{t-} \nu_2(dh_j)} \\ &= \frac{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2} e^{-\bar{\nu}_1 t} \prod_{t-} \nu_1(dh_j)}{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2} e^{-\bar{\nu}_1 t} \prod_{t-} \nu_1(dh_j) + (1 - q_0) e^{\mu_2 Y_B^1(t)/\sigma^2 - \mu_2^2 t/2\sigma^2} e^{-\bar{\nu}_2 t} \prod_{t-} \nu_2(dh_j)}. \end{aligned}$$

Indeed, $\frac{1}{\sqrt{2\pi t \sigma}} e^{-\frac{(Y_B^1(t) - \mu_1 t)^2}{2\sigma^2 t}}$ is the probability of receiving the payoff $Y_B^1(t)$, given the type θ_1 , and $e^{-\bar{\nu}_1 t} \frac{(\bar{\nu}_1 t)^n}{n!} \prod_{t-} \frac{\nu_1(dh_j)}{\bar{\nu}_1}$ is the probability of receiving the n jumps that occurred until time t , given the type θ_1 . The first equality in (2.4) is the Bayesian belief updating, using the independence of the components in the Lévy-Ito decomposition, given the type of the risky arm, and the second equality is obtained by eliminating common components. For a generic path of the process (q_t) , see figure 3.

Suppose the DM follows a cut-off strategy κ' with cut-off point p' :

$$\kappa' = \begin{cases} 0 & \text{if } p \leq p', \\ 1 & \text{if } p > p'. \end{cases}$$

If $q_0 \leq p'$, the DM will always choose the safe arm. If $q_0 > p'$, the DM will initially choose the risky arm. The DM chooses the risky arm as long as $q_t > p'$, which, by Eq. (2.4), is equivalent to:

$$(2.5) \quad \frac{q_0 e^{\mu_1 Y_B^1(t)/\sigma^2 - \mu_1^2 t/2\sigma^2 - \bar{\nu}_1 t}}{(1 - q_0) e^{\mu_2 Y_B^1(t)/\sigma^2 - \mu_2^2 t/2\sigma^2 - \bar{\nu}_2 t}} \prod_{t-} \frac{\nu_1(dh)}{\nu_2(dh)} > \frac{p'}{1 - p'}.$$

Without loss of generality assume that $\mu_1 - \mu_2 > 0$; the case $\mu_1 - \mu_2 < 0$ is handled similarly, and provides the same results. By taking the natural logarithm, and rearranging the resulting terms, we obtain that this inequality is equivalent to:

$$(2.6) \quad \begin{aligned} \frac{1}{\sigma} Y_B^1(t) &> \left[\left(\frac{\mu_1 + \mu_2}{2\sigma} \right) + \frac{\sigma(\bar{\nu}_1 - \bar{\nu}_2)}{\mu_1 - \mu_2} \right] t - \frac{\sigma}{\mu_1 - \mu_2} \times \left[\ln \left(\frac{q_0}{1 - q_0} \right) - \ln \left(\frac{p'}{1 - p'} \right) \right] \\ &\quad - \frac{\sigma}{\mu_1 - \mu_2} \sum_{t-} \ln \left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)} \right). \end{aligned}$$

The right-hand side in (2.6) is a piecewise linear function $F \cdot t - E - G_t$ of t , where the slope $F := \left(\frac{\mu_1 + \mu_2}{2\sigma} \right) + \frac{\sigma(\bar{\nu}_1 - \bar{\nu}_2)}{\mu_1 - \mu_2}$ is independent of t , the intercept at $t = 0$ is $E := \frac{\sigma}{\mu_1 - \mu_2} \times \left[\ln \left(\frac{q_0}{1 - q_0} \right) - \ln \left(\frac{p'}{1 - p'} \right) \right]$, and $G_t := \frac{\sigma}{\mu_1 - \mu_2} \sum_{t-} \ln \left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)} \right)$. Denote by $G_h := \frac{\sigma}{\mu_1 - \mu_2} \ln \left(\frac{\nu_1(dh)}{\nu_2(dh)} \right)$ the contribution of a jump of size h to the intercept.

⁸ $\prod_{t-} \nu_i(dh_j)$ is the product of $\nu_i(dh_j)$ over all jumps h_j that occur up to time $t-$. Similarly, we use the notation \sum_{t-} for the sum over all jumps up to time $t-$.

From Eq. (2.6) we obtain the following alternative description of the optimal strategy: The DM has a time-dependent cut-off which is piecewise linear. The slope of the cut-off function is always F , and whenever there is a jump of size h , the cut-off decreases by G_h (see Figure 4). The DM chooses the risky arm as long as his current payoff from the continuous part of the Lévy process, $Y_B^1(t)$, exceeds the cut-off.

Thus, at first the DM plays until the payoff from the continuous part divided by the standard deviation satisfies: $\frac{1}{\sigma} Y_B^1(t) \leq F \cdot t - E$; if a jump of size h occurs before he switches to the safe arm, then the intercept decreases by G_h , and this behavior repeats itself. If there is no Brownian motion component, then $Y_B^1(t) \equiv 0$ and $F = 0$; the DM chooses the risky arm for a fixed amount of time, and then switches to the safe arm, unless a jump occurs before the switch; if a jump occurred, the amount of time to choose the risky arm increases, as a function of $\frac{\nu_1(dh)}{\nu_2(dh)}$.

Theorem 2.5. *Under Assumption 2.1, for every $p_0 \in [0, 1]$, the payoff $U(p_0, q_0)$ of a DM who uses a cut-off strategy with cut-off point p' , is as follows:⁹ if $q_0 \geq p'$,*

$$(2.7) \quad U(p_0, q_0) = (s - g_1)p_0 \left(\frac{1 - q_0}{q_0} \right)^{\alpha+1} \left(\frac{p'}{1 - p'} \right)^{\alpha+1} + (s - g_2)(1 - p_0) \left(\frac{1 - q_0}{q_0} \right)^{\alpha} \left(\frac{p'}{1 - p'} \right)^{\alpha} \\ + g_1 p_0 + g_2(1 - p_0),$$

while $U(p_0, q_0) = s$ if $q_0 < p'$.

One can verify that when the DM holds the correct prior, and plays according to the optimal strategy κ^* , Eq. (2.7) coincides with (2.3): $U(p_0, p_0) = U(p_0)$. Note that $U(p_0, q_0)$ is continuous in all its parameters.

To calculate the payoff of a DM who plays optimally but has an incorrect prior, simply substitute the expression for p^* from Theorem 2.3 as p' in Theorem 2.5.

2.5. Information and Payoff. Theorems 2.3 and 2.5 express the optimal cut-off, the optimal expected payoff, and the expected payoff for a DM with incorrect prior who uses a cut-off strategy in terms of the expected payoff of each arm, and the quantity α , which captures all the information-relevant parameters of the payoff processes. In this section we justify why α indeed captures all the information-relevant parameters. This justification explains why the solutions of Bolton and Harris (1995) and Keller, Rady and Cripps (2005) have the same structure as our solution. Suppose the DM faces a two-armed bandit problem with Lévy payoffs. If the risky arm's type is High (resp. Low), it yields three independent Lévy processes $(X_1^a(t))$, $(X_1^b(t))$, $(X_1^c(t))$ (resp. $(X_2^a(t))$, $(X_2^b(t))$, $(X_2^c(t))$).

As in Section 2.2, the notation μ_i^j , σ_i^j , ν_i^j , and g_i^j , represents the drift of the process $(X_i^j(t))$, the standard deviation of the Brownian motion component of $(X_i^j(t))$, the Lévy measure of the process $(X_i^j(t))$, and the expectation of the process $(X_i^j(t))$ at time $t = 1$, respectively, where $j \in \{a, b, c\}$, and $i \in \{1, 2\}$ is the arm's type.

We assume that Assumption 2.1 is satisfied for the three couples $((X_1^j(t)), (X_2^j(t)))$, $j \in \{a, b, c\}$.

⁹We omit the dependence of U on p' , s , g_1 , g_2 . Recall that α is the unique solution of the equation $f(\eta) = 0$ (see Eq. (4.4)).

Suppose that the DM's payoff is $X_i^a + X_i^c$, while his information is (X_i^a, X_i^b) . Thus, the component X_i^a represents observed information of the DM that contributes to his payoff, the component X_i^b represents observed information that does not contribute to the payoff, and the component X_i^c represents unobserved information that contributes to the payoff. The following theorem characterizes the optimal strategy of the DM in this setup, as well as his expected payoff when he holds a possibly incorrect prior. This characterization can be used to calculate the fair price of additional information.

Let β be the unique solution of

$$(2.8) \quad \begin{aligned} f_{a,b}(\eta) := & \int \nu_2^a(dh^a) \left(\frac{\nu_2^a(dh^a)}{\nu_1^a(dh^a)} \right)^\eta + \eta(\bar{\nu}_1^a - \bar{\nu}_2^a) - \bar{\nu}_2^a + \frac{1}{2}(\eta+1)\eta \left(\frac{\mu_1^a - \mu_2^a}{\sigma^a} \right)^2 \\ & + \int \nu_2^b(dh^b) \left(\frac{\nu_2^b(dh^b)}{\nu_1^b(dh^b)} \right)^\eta + \eta(\bar{\nu}_1^b - \bar{\nu}_2^b) - \bar{\nu}_2^b + \frac{1}{2}(\eta+1)\eta \left(\frac{\mu_1^b - \mu_2^b}{\sigma^b} \right)^2 - r = 0 \end{aligned}$$

in $(0, \infty)$.

Theorem 2.6. *The expected payoff to a DM who holds the prior belief q_0 , uses a cut-off strategy κ' with cut-off point p' , receives and observes the payoff process $(X^a(t))$, observes but does not receive the payoff process $(X^b(t))$, and receives but does not observe the payoff process $(X^c(t))$, is as follows: if $q_0 > p'$*

$$\begin{aligned} V_{p'}^{a,b,c}(p_0, q_0) := & (s - g_1^a - g_1^c)p_0 \left(\frac{1 - q_0}{q_0} \right)^{\beta+1} \left(\frac{p'}{1 - p'} \right)^{\beta+1} \\ & + (s - g_2^a - g_2^c)(1 - p_0) \left(\frac{1 - q_0}{q_0} \right)^\beta \left(\frac{p'}{1 - p'} \right)^\beta \\ & + p_0(g_1^a + g_1^c) + (1 - p_0)(g_2^a + g_2^c), \end{aligned}$$

while $V_{p'}^{a,b,c}(p_0, q_0) = s$ if $q_0 \leq p'$.

Moreover, the optimal cut-off for a DM who holds the correct prior, i.e. $q_0 = p_0$, is given by

$$p^* = \frac{\beta(s - g_2^a - g_2^c)}{(\beta + 1)(g_1^a + g_1^c - s) + \beta(s - g_2^a - g_2^c)}.$$

Observe that if there are no processes (X_i^b, X_i^c) , then Theorem 2.6 reduces to Theorem 2.5. Moreover, the parameter β in Theorem 2.6 is the equivalent of α in Theorem 2.5.

The characterization in Theorem 2.6 shows that β incorporates all the data that is relevant to the information that the DM has. It depends on the parameters of (X_i^a, X_i^b) , which the DM observes, but not on the parameters of (X_i^c) , which the DM does not observe. Moreover, it depends on the absolute value of the difference between the drifts, $|\mu_1 - \mu_2|$, the standard deviation of the Brownian motion, σ , the Radon-Nikodym derivative $\frac{\nu_2(dh)}{\nu_1(dh)}$, and the average $\eta\bar{\nu}_1 + (1 - \eta)\bar{\nu}_2$. These quantities help in distinguishing between the two types of the Risky arm. Once the information-relevant parameters are summarized in β , the only relevant parameters, which affect both the optimal cut-off and the optimal expected payoff, are the expectations of the payoff-relevant processes (g_i^a, g_i^c) .

Theorem 2.6 can be used to find the fair price of the additional information X^b and of X^c . That is, the price of information that does not affect the DM's payoff, as well as the price of information that affects the DM's payoff. This is done by comparing the optimal value of two bandit problems that differ in the information of the DM.

The next corollary to Theorem 2.6 states that, using cut-off strategies, additional information is more profitable.¹⁰

Corollary 2.7. *Suppose that there are two decision makers, DM1 and DM2, who hold the correct prior, i.e. $q_0 = p_0$, and use cut-off strategies. DM1 receives and observes the payoff process $(X^a(t))$, observes but does not receive the payoff process $(X^b(t))$, and receives but does not observe the payoff process $(X^c(t))$. DM2 receives and observes the payoff process $(X^a(t))$, and receives but does not observe the payoff process $(X^c(t))$. Then the optimal expected payoff of DM1 is higher than the optimal expected payoff of DM2.*

2.6. Optimism vs Pessimism. As mentioned before, the phenomenon of over confidence is common in many decision problems. In this section we apply the result of Section 2.4, and investigate who will fare better in two-armed bandit problems, an optimist who assigns a probability higher than the true probability to the High type, or a pessimist who assigns a probability lower than the true probability to the High type.

Suppose that there are two decision makers, DM1 and DM2, who face independent identical copies of the decision problem. DM1 is an *optimist*, and believes that the probability of High is $p_0 + \rho$, where $\rho > 0$, while DM2 is a *pessimist*, and believes that the probability of High is $p_0 - \rho$, and both play optimally given their beliefs. If $p_0 - \rho \leq p^*$, the pessimist will always choose the safe arm, since according to his subjective belief the prior is at most the cut-off. Assume then that $p_0 - \rho > p^*$. For every $\epsilon \in [p^* - p_0, 1 - p_0]$ denote by $V_{p_0}(\epsilon)$ the expected payoff for a DM playing optimally according to the incorrect prior $p_0 + \epsilon$, where ϵ may be negative. It turns out that the answer regarding who will fare better, an optimist or a pessimist, depends on α that is defined in Eq. (2.2).

Theorem 2.8. *Assume that Assumption 2.1 holds.*

1. *If $\alpha > 1$, then for every $\epsilon > 0$ such that $p^* < p_0 \pm \epsilon \leq 1$ we have $V_{p_0}(\epsilon) > V_{p_0}(-\epsilon)$: an optimist will fare better.*
2. *If $0 < \alpha < 1$, then for every $\epsilon > 0$ such that $p^* < p_0 \pm \epsilon \leq \frac{\alpha+2}{3}$ we have $V_{p_0}(\epsilon) > V_{p_0}(-\epsilon)$: an optimist will fare better; for every $\epsilon > 0$ such that $\frac{\alpha+2}{3} < p_0 \pm \epsilon < 1$ we have $V_{p_0}(\epsilon) < V_{p_0}(-\epsilon)$: a pessimist will fare better.*

Thus, an optimist will fare better than a pessimist, unless α is low and $p_0 - \epsilon$ is high. That is, a pessimist will fare better only if the following two conditions are satisfied:

- 1) The pessimist assigns high probability to the High type.
- 2) The two DM's are sufficiently patient, or it is easy to distinguish between the

¹⁰It is well known that in one-player optimization problems, additional information cannot hurt the DM, since he can always ignore it. However, a Markovian cut-off strategy does not allow a player to forget additional information, and therefore the statement is not trivial.

two types of the risky arm.
Otherwise, an optimist will fare better.

Since $p^* = \frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$, the condition $p^* < \frac{\alpha+2}{3}$ is not always satisfied. If $3g_1 + g_2 \geq 4s$, then $p^* < \frac{\alpha+2}{3}$. If $3g_1 + g_2 < 4s$, then $p^* < \frac{\alpha+2}{3}$ if and only if $\alpha < \frac{4s-3g_1-g_2-\sqrt{(4s-3g_1-g_2)^2-8(g_1-g_2)(g_1-s)}}{2(g_1-g_2)}$ or $\alpha > \frac{4s-3g_1-g_2+\sqrt{(4s-3g_1-g_2)^2-8(g_1-g_2)(g_1-s)}}{2(g_1-g_2)}$.

We obtain the same result if the biases of the optimist and the pessimist are geometric rather than absolute, that is, if the initial prior of the optimist is $(1+\epsilon)p_0 = p_0 + \epsilon p_0$ and the initial prior of the pessimist is $(1-\epsilon)p_0 = p_0 - \epsilon p_0$. This is true since the absolute bias of the optimist and the pessimist is the same: ϵp_0 .

3. FUTURE DIRECTIONS

Our results call for further research. We here list few possible directions for future research.

- We studied the case that the distribution of the High type dominates the distribution of the Low type in a strong sense (see Assumptions 2.1). These assumptions ensure that the discontinuities of the process of posterior belief are always to one direction. It would be interesting to know whether a similar characterization holds under only assumptions A1 and A2.
- We assumed that the payoff is distributed according to a Lévy process. It would be interesting to solve the model when the payoff is distributed according to a geometric Lévy process.
- Bolton and Harris (1999), Keller, Rady and Cripps (2004), and Keller and Rady (2008) study a strategic version of the model, in which several decision makers face identical unknown arms. Klein and Rady (2008) study a strategic version of the model, in which the arms of two DMs are negatively correlated. It will be interesting to solve these models, when the payoff distributions are general Lévy processes.

4. PROOFS

4.1. Proof of Proposition 2.2. We already argued that U is non-decreasing and convex. It is left to prove that it is continuous. Since $U(p)$ is convex, it is continuous on $(0, 1)$. If the DM had known the true type of the arm, his optimal payoff would have been $pg_1 + (1-p)s$. Since this information is not available, $U(p) \leq pg_1 + (1-p)s$. Since $U(p) \geq s$ for every p , this implies $\lim_{p \rightarrow 0} U(p) = s = U(0)$. Since the DM can follow the strategy that always selects the risky arm, $U(p) \geq pg_1 + (1-p)g_2$. This implies that $\lim_{p \rightarrow 1} U(p) = g_1 = U(1)$. Thus, U is continuous on $[0, 1]$.

4.2. The Control Equation. In this section we express the optimal payoff $U(p)$ using the dynamic programming principle. This representation extends that in Bolton and Harris (1999) to the more general setup of Lévy payoff processes. We start by calculating how the belief of the DM is updated given his observations. To simplify notation, it is convenient to divide various expressions by the standard deviation σ ; set $\tilde{\mu}_1 = \frac{\mu_1}{\sigma}$, $\tilde{\mu}_2 = \frac{\mu_2}{\sigma}$ and $\tilde{\mu} = \frac{\mu}{\sigma}$, and denote $d\tilde{Y}_B^k := \frac{1}{\sqrt{k}\sigma} dY_B^k =$

$\sqrt{k}\tilde{\mu}dt + dZ(t)$, modulo the factor $\sqrt{\kappa}\sigma$, $d\tilde{Y}_B^k$ is the contribution to the payoff of the linear drift and the Brownian motion. Then the probability of obtaining a specific observation (that is not a jump) given the type is:

$$\begin{aligned} P(d\tilde{Y}_B^k|\tilde{\mu}) &= P(d\tilde{Y}_B^k|\mu) = \frac{1}{\sqrt{2\pi dt}} e^{-\frac{(d\tilde{Y}_B^k(t) - \sqrt{k}\tilde{\mu}dt)^2}{2dt}} = C \cdot e^{\sqrt{k}\tilde{\mu}d\tilde{Y}_B^k - \frac{k\tilde{\mu}^2(dt)^2}{2}} \\ &= C(1 + \sqrt{k}\tilde{\mu}d\tilde{Y}_B^k + o((dt)^2)), \end{aligned}$$

where C is a constant independent of μ . Denote $\tilde{P}(d\tilde{Y}_B^k|\tilde{\mu}) := 1 + \sqrt{k}\tilde{\mu}d\tilde{Y}_B^k$.

Let $p = P_t(\theta_1)$ be the belief at time t that the risky arm is High. Then:

$$\begin{aligned} P_{t+dt}(\theta_1) &= \frac{P(dY_B^k|\theta_1)P(dY_P^k|\theta_1)p}{P(dY_B^k|\theta_1)P(dY_P^k|\theta_1)p + P(dY_B^k|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\ &= \frac{P(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1)p}{P(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1)p + P(d\tilde{Y}_B^k|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\ &= \frac{\tilde{P}(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1)p}{\tilde{P}(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1)p + \tilde{P}(d\tilde{Y}_B^k|\theta_2)P(dY_P^k|\theta_2)(1-p)}. \end{aligned}$$

This is the Bayesian belief updating, using the independence of the components in the Lévy-Ito decomposition, given the risky-arm type θ . The next lemma expresses the change in the posterior belief over time. The lemma handles separately the case where there are no jumps in the time interval $[t, t+dt)$, and the case where there is a jump in this interval.

Lemma 4.1.

1. Suppose that there are no jumps during the time interval $[t, t+dt)$. Then:

$$(4.1) \quad dP_t := P_{t+dt}(\theta_1) - P_t(\theta_1) = p(1-p)\sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)d\tilde{Z} - p(1-p)k(\bar{\nu}_1 - \bar{\nu}_2)dt,$$

where $d\tilde{Z} = d\tilde{Y}_B^k - \sqrt{k}(p\tilde{\mu}_1 + (1-p)\tilde{\mu}_2)dt$ is a standard Brownian motion.

2. Suppose that during the interval $[t, t+dt)$ a jump of size h occurred. Then:

$$(4.2) \quad P_{t+dt}(\theta_1) = P_h + \sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)P_h(1-P_h)d\tilde{Z}_2,$$

where $P_h := \frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)}$, and $d\tilde{Z}_2 := d\tilde{Y}_B^k - \sqrt{k}(\tilde{\mu}_1 P_h + \tilde{\mu}_2(1-P_h))dt$.

The first term in the right-hand side of (4.2) is the contribution of the continuous part of the payoff process to the change in the belief, while the second term is the contribution of the fact that no jump arrived. This latter contribution is negative due to Assumption A3. If a jump of size h occurs during the interval $[t, t+dt)$, then the contribution of the continuous part of the payoff process is $(\tilde{\mu}_1 - \tilde{\mu}_2)(d\tilde{Y}_B^k - \sqrt{k}(\tilde{\mu}_1 P_h + \tilde{\mu}_2(1-P_h))dt)$, and the compound Poisson process' contribution is $P_h := \frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)}$. The latter is the Bayesian update of the probability that the risky arm is High given that a jump of size h occurred. Note that $p < P_h$.

Proof of Lemma 4.1. The proof of the lemma is standard and non-inspiring. The first statement follows from a long chain of equalities. Assume that there is no

jump in the interval $[t, t + dt)$

$$\begin{aligned}
dP_t &= \frac{p(1-p)[\tilde{P}(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1) - \tilde{P}(d\tilde{Y}_B^k|\theta_2)P(dY_P^k|\theta_2)]}{\tilde{P}(d\tilde{Y}_B^k|\theta_1)P(dY_P^k|\theta_1)p + \tilde{P}(d\tilde{Y}_B^k|\theta_2)P(dY_P^k|\theta_2)(1-p)} \\
&= \frac{p(1-p)[(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)(e^{-\bar{\nu}_1 k dt}) - (1 + \sqrt{k}\tilde{\mu}_2 d\tilde{Y}_B^k)(e^{-\bar{\nu}_2 k dt})]}{(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)(e^{-\bar{\nu}_1 k dt})p + (1 + \sqrt{k}\tilde{\mu}_2 d\tilde{Y}_B^k)(e^{-\bar{\nu}_2 k dt})(1-p)} \\
&= \frac{p(1-p)[(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)(1 - \bar{\nu}_1 k dt) - (1 + \sqrt{k}\tilde{\mu}_2 d\tilde{Y}_B^k)(1 - \bar{\nu}_2 k dt)]}{(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)(1 - \bar{\nu}_1 k dt)p + (1 + \sqrt{k}\tilde{\mu}_2 d\tilde{Y}_B^k)(1 - \bar{\nu}_2 k dt)(1-p)} \\
&= \frac{p(1-p)[\sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)d\tilde{Y}_B^k - (\bar{\nu}_1 - \bar{\nu}_2)k dt]}{1 + \sqrt{k}(p\tilde{\mu}_1 + (1-p)\tilde{\mu}_2)d\tilde{Y}_B^k - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)dt} \\
&= p(1-p)[\sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)d\tilde{Y}_B^k - (\bar{\nu}_1 - \bar{\nu}_2)k dt] \\
&\quad \cdot [1 - \sqrt{k}(p\tilde{\mu}_1 + (1-p)\tilde{\mu}_2)d\tilde{Y}_B^k + k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)dt + k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)^2 dt] \\
&= p(1-p)[\sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)(d\tilde{Y}_B^k - \sqrt{k}(p\tilde{\mu}_1 + (1-p)\tilde{\mu}_2)dt) - k(\bar{\nu}_1 - \bar{\nu}_2)dt] \\
&= p(1-p)\sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)d\tilde{Z} - p(1-p)k(\bar{\nu}_1 - \bar{\nu}_2)dt.
\end{aligned}$$

In the calculations we used the fact that $d\tilde{Z} = d\tilde{Y}_B^k - \sqrt{k}(p\tilde{\mu}_1 + (1-p)\tilde{\mu}_2)dt$ is a standard Brownian motion (see Bolton and Harris (1999)), and the Brownian motion properties: $dZ^2 = dt$, and $dZdt = 0$. We also ignored terms of order $(dt)^{3/2}$ and above.

We now prove the second statement.

$$\begin{aligned}
P_{t+dt}(\theta_1) &= \frac{p\nu_1(dh)(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)}{[\nu_1(dh)p + \nu_2(dh)(1-p)] + \sqrt{k}[\nu_1(dh)\tilde{\mu}_1 p + \nu_2(dh)\tilde{\mu}_2(1-p)]d\tilde{Y}_B^k} \\
&= \frac{P_h(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)}{1 + \sqrt{k}[\tilde{\mu}_1 P_h + \tilde{\mu}_2(1-P_h)]d\tilde{Y}_B^k} \\
&= P_h(1 + \sqrt{k}\tilde{\mu}_1 d\tilde{Y}_B^k)[1 - \sqrt{k}(P_h\tilde{\mu}_1 + (1-P_h)\tilde{\mu}_2)d\tilde{Y}_B^k + k(P_h\tilde{\mu}_1 + (1-P_h)\tilde{\mu}_2)^2 dt] \\
&= P_h[1 + \sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)(1-P_h)(d\tilde{Y}_B^k - \sqrt{k}(\tilde{\mu}_1 P_h + \tilde{\mu}_2(1-P_h))dt)] \\
&= P_h + \sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)P_h(1-P_h)d\tilde{Z}_2.
\end{aligned}$$

□

We now formulate the control equation that describes the optimal payoff.

$$\begin{aligned}
&(\text{CE}) \\
U(p) &= \max_{k \in [0,1]} \{[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]r dt + e^{-r dt} E[U(p+dp)]\},
\end{aligned}$$

where k is the control variable. The first term within the maximization is the expected instantaneous payoff, and the second term is the discounted expected

continuation payoff. The following lemma provides a more convenient form to the control equation, in terms of the derivatives of U .

Lemma 4.2. *The following equality holds:*

(CE2)

$$\begin{aligned} U(p) = \max_{k \in [0,1]} \{ & (1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2)) \\ & + \frac{1}{r} \left[k \int (p\nu_1(dh) + (1-p)\nu_2(dh)) U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) \right. \\ & - kp(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p) - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p) \\ & \left. + \frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2 \right] \}, \quad p - a.s. \end{aligned}$$

Proof of Lemma 4.2. Since $U(p)$ is convex, $U(p)$ is twice differentiable p -a.s. (in the sense of the Lebesgue measure). With probability $[p(1 - k\bar{\nu}_1 dt) + (1-p)(1 - k\bar{\nu}_2 dt)]$ there are no jumps in the interval $[t, t + dt]$. In this case, using the Taylor expansion of U , and ignoring terms of order $(dt)^{3/2}$ and higher, we obtain that the optimal payoff is $U(p + \check{d}p) = U(p) + U'(p)\check{d}p + \frac{1}{2}U''(p)(\check{d}p)^2$ a.s.,¹¹ where $\check{d}p$ is given by the right-hand side of (4.1).

With probability $[pk\nu_1(dh)dt + (1-p)k\nu_2(dh)dt]$ there is a jump of size h , and the optimal payoff is $U(P_h + \sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)P_h(1 - P_h)d\tilde{Z}_2) = U(P_h) + U'(P_h)\hat{d}P_h + \frac{1}{2}U''(P_h)\hat{d}P_h$, where $\hat{d}P_h := \sqrt{k}(\tilde{\mu}_1 - \tilde{\mu}_2)P_h(1 - P_h)d\tilde{Z}_2$.

During the subsequent calculations we use the following Eqs. (4.3), (4.4), (4.5) and (4.6), that can be derived from Lemma 4.1:

$$(4.3) \quad E[\check{d}p] = -kp(1-p)(\bar{\nu}_1 - \bar{\nu}_2)dt.$$

This is the expectation of the change in the belief, given that no jump occurred during the interval $[t, t + dt]$.

$$(4.4) \quad E[\check{d}p^2] = kp^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2dt.$$

This is the second moment of the change in the belief, given that no jump occurred during the interval $[t, t + dt]$.

$$(4.5) \quad E[\hat{d}P_h] = C_1 \cdot dt.$$

This is the expected contribution of the Brownian motion part to the posterior belief, given that a jump of size h occurred during the interval $[t, t + dt]$.

$$(4.6) \quad E[\hat{d}P_h^2] = C_2 \cdot dt.$$

This is the second moment of the contribution of the Brownian motion part to the posterior belief, given that a jump of size h occurred during the interval $[t, t + dt]$. In Eqs. (4.5) and (4.6), C_1 and C_2 are constants. Using the above notation, we

¹¹Since we can ignore terms of order $(dt)^{3/2}$ and higher, it is sufficient to consider the Taylor expansion up to the second derivative.

obtain from (CE):

(4.7)

$$\begin{aligned}
U(p) = & \max_{k \in [0,1]} \{[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]r dt \\
& + (1-r dt) \left[k dt \left[\int \left[U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) \right. \right. \right. \\
& \left. \left. \left. + U' \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) C_1 dt \right. \right. \right. \\
& \left. \left. \left. + U'' \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) C_2 dt \right] (p\nu_1(dh) + (1-p)\nu_2(dh)) \right] \right. \\
& \left. + [1 - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)dt] \left[U(p) + U'(p)E[\check{d}p] + \frac{1}{2}U''(p)E[\check{d}p^2] \right] \right\}, \quad p - a.s.
\end{aligned}$$

The second, third and fourth lines in (4.7) represent the expected continuation payoff given a jump of size h occurred during the time interval $[t, t + dt)$, and the fifth line represents the expected continuation payoff given no jump occurred during the time interval $[t, t + dt)$. Using $(dt)^2 = 0$, several of the terms in (4.7) vanish, and we obtain:

(4.8)

$$\begin{aligned}
U(p) = & \max_{k \in [0,1]} \{[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]r dt \\
& + k dt \int U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) (p\nu_1(dh) + (1-p)\nu_2(dh)) \\
& + U(p) - k dt p(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p)dt - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p)dt - rU(p)dt \\
& + \frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2 dt \Big\}, \quad p - a.s.
\end{aligned}$$

Eliminating $U(p)$ from both sides, and dividing by dt , we obtain (CE2) after simple algebraic manipulations, as desired. \square

From Eq. (4.7) it follows that the contribution of the continuation payoff given that a jump of size h occurred during the time interval $[t, t + dt)$ is

(4.9)

$$\begin{aligned}
& k dt \left[\int \left[U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) + U' \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) C_1 dt \right. \right. \\
& \left. \left. + U'' \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) C_2 dt \right] (p\nu_1(dh) + (1-p)\nu_2(dh)) \right].
\end{aligned}$$

The parameters of the Brownian motion affect (4.9) only through C_1 and C_2 , and since C_1 and C_2 do not appear in Eq. (4.8), it follows that if a jump occurs during the time interval $[t, t + dt)$, the information from the compound Poisson process has more impact than the information of the Brownian motion.

According to (CE2), the payoff is the maximum over the control variable k of the expectation of the current flow payoff $[(1-k)s + k(p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2))]$

plus the discounted value of the continuation payoff

$$\frac{1}{r} \left[k \int (p\nu_1(dh) + (1-p)\nu_2(dh)) U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) - kp(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p) + \frac{1}{2}kU''(p)p^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2 - k(p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p) \right].$$

A solution κ to this maximization problem must satisfy:

$$(4.10) \quad k = \begin{cases} 0 & \text{if } b(p, U) < s - [pg_1 + (1-p)g_2], \\ \in [0, 1] & \text{if } b(p, U) = s - [pg_1 + (1-p)g_2], \\ 1 & \text{if } b(p, U) > s - [pg_1 + (1-p)g_2], \end{cases}$$

where

$$b(p, U) = \frac{1}{r} \left[\int (p\nu_1(dh) + (1-p)\nu_2(dh)) U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) - p(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p) - (p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p) + \frac{1}{2}U''(p)p^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2 \right].$$

The function within the maximization in (CE2) is linear in k . Therefore, it achieves its maximum at $k = 1$ or $k = 0$, p -a.s. From Proposition 2.2, $U(p)$ is non-decreasing and continuous, and therefore there is p^* such that $U(p) = s$ for every $p \leq p^*$. Thus $k = 0$ is optimal for $p < p^*$. For every $p > p^*$, we have $U(p) > s$, so that in this case $k = 1$ is optimal p -a.s.¹²

4.3. Characterizing the optimal strategy and the value. When it is optimal to play safe, that is, when the optimal solution of (CE) is $k^* = 0$, we have $U(p) = s$. When it is optimal to play risky, that is, when the optimal solution of (CE) is $k^* = 1$, it follows from Lemma 4.2 that $U(p)$ solves the following functional differential equation:

$$(FDE) \quad U(p) = p(\bar{\nu}_1 H_1 + \mu_1) + (1-p)(\bar{\nu}_2 H_2 + \mu_2) + \frac{1}{r} \left[\int (p\nu_1(dh) + (1-p)\nu_2(dh)) U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) - p(1-p)(\bar{\nu}_1 - \bar{\nu}_2)U'(p) - (p\bar{\nu}_1 + (1-p)\bar{\nu}_2)U(p) + \frac{1}{2}U''(p)p^2(1-p)^2(\tilde{\mu}_1 - \tilde{\mu}_2)^2 \right], \text{ a.s. in } (p^*, 1).$$

A solution $U(p)$ for this equation must be smooth (Friedman (1969), p.56).¹³ Therefore, $U(p)$ satisfies Eq. (FDE) in $(p^*, 1)$ always, and $k = 1$ is optimal in $(p^*, 1)$.

¹²Recall that Eq. (CE2) is satisfied p -a.s., since $U'(p)$ and $U''(p)$ exist p -a.s. In the next Section we show that the optimal strategy is in fact a cut-off strategy.

¹³To see that the conditions of Friedman (1969) are satisfied, substitute $f(p) = \int (p\nu_1(dh) + (1-p)\nu_2(dh)) U \left(\frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \right) - U(p)(r + (p\bar{\nu}_1 + (1-p)\bar{\nu}_2))$. Since $U(p)$ is continuous, so is f , and we get $U \in C^2$, as Friedman (1969) requires.

Hence, there is an optimal cut-off strategy κ^* with cut-off point p^* .

The next lemma suggests one solution to Eq. (FDE). To prove it, substitute the expression for $U(p)$ defined in Eq. (4.11) below into Eq. (FDE). Recall that

$$f(\eta) = \int \nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\bar{\nu}_1 - \bar{\nu}_2) - \bar{\nu}_2 + \frac{1}{2}(\eta + 1)\eta(\tilde{\mu}_1 - \tilde{\mu}_2)^2 - r$$

(see (2.2)), that α is the unique solution of the equation $f(\eta) = 0$ in $(0, \infty)$, and that p^* and C_α were defined in the statement of Theorem 2.3.

Lemma 4.3. *One smooth solution to Eq. (FDE) is*

$$(4.11) \quad U(p) = pg_1 + (1-p)g_2 + C_\alpha(1-p) \left(\frac{1-p}{p} \right)^\alpha,$$

where $\alpha \in (0, \infty)$ solves the equation $f(\eta) = 0$.

In order to see that the function $U(p)$, defined above, actually solves (FDE), we use the fact that $P_h = \frac{p\nu_1(dh)}{p\nu_1(dh) + (1-p)\nu_2(dh)} \geq p^*$ for every $p > p^*$, ν_1 -a.s., which is equivalent to Assumption A3. Thus, the form of our solution crucially depends on this assumption.

In fact, one can verify that

$$U(p) = pg_1 + (1-p)g_2 + C(1-p) \left(\frac{1-p}{p} \right)^\alpha + D(1-p) \left(\frac{1-p}{p} \right)^\beta$$

solves Eq. (FDE), where α is as in the statement of Lemma 4.3, and β is the unique solution¹⁴ of $f(\eta) = 0$ in $(-\infty, 0)$. The following lemma assures that α is well defined.

Lemma 4.4. *The equation $f(\eta) = 0$ admits a unique solution in the interval $(0, \infty)$.*

Proof.

The function f is a continuous function that satisfies $f(0) < 0$ and $f(\infty) = \infty$. To show that $f(\eta) = 0$ has a unique solution, it is therefore sufficient to prove that f is increasing in η . Note that if $\mu_1 \neq \mu_2$, then $\frac{1}{2}(\eta + 1)\eta(\tilde{\mu}_1 - \tilde{\mu}_2)^2 - r - \bar{\nu}_2$ is increasing in η , and constant otherwise. It remains to prove that if $\nu_1 \neq \nu_2$ (i.e. $\nu_1(\mathbb{R} \setminus \{0\}) > \nu_2(\mathbb{R} \setminus \{0\})$), then $\int \nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\bar{\nu}_1 - \bar{\nu}_2)$ is increasing in η . Since

$$\int \nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\bar{\nu}_1 - \bar{\nu}_2) = \int \left[\nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\nu_1(dh) - \nu_2(dh)) \right],$$

and

$$\int_{\{h | \frac{\nu_2(dh)}{\nu_1(dh)} = 1\}} \left[\nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\nu_1(dh) - \nu_2(dh)) \right] = 0,$$

it is sufficient to prove that for ν_1 -a.e. $h \in \{h | \frac{\nu_2(dh)}{\nu_1(dh)} < 1\}$,

$$g_h(\eta) = \nu_2(dh) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + \eta(\nu_1(dh) - \nu_2(dh))$$

¹⁴In fact, such a solution β must be smaller than -1 .

is increasing in η . Now,

$$\begin{aligned} g'_h(\eta) &= -\nu_2(dh) \ln \left(\frac{\nu_1(dh)}{\nu_2(dh)} \right) \left(\frac{\nu_2(dh)}{\nu_1(dh)} \right)^\eta + (\nu_1(dh) - \nu_2(dh)) \\ &> -\nu_2(dh) \ln \left(\frac{\nu_1(dh)}{\nu_2(dh)} \right) + (\nu_1(dh) - \nu_2(dh)) > 0, \end{aligned}$$

where the first inequality holds since $\eta > 0$ and by Assumption A3. The second inequality holds since $-\ln(x) + x - 1 > 0$ for every $x \neq 1$. Therefore, $g_h(\eta)$ is increasing, as desired. \square

We now prove that Eq. (FDE) has a unique solution.

Lemma 4.5. *For every $p_1 < p_2$, and every $u_1, u_2 \in \mathbb{R}$, there is a unique solution $U(p)$ satisfying Eq. (FDE) in the interval (p_1, p_2) with the boundary conditions $U(p_1) = u_1$, $U(p_2) = u_2$.*

Proof.

Since Eq. (FDE) is a non-homogenous linear equation in U , if there are two solutions of Eq. (FDE), then their difference is a solution of the homogenous version of Eq. (FDE). To prove the lemma, it is therefore sufficient to fix a solution W of the homogenous version of Eq. (FDE) that satisfies $W(p_1) = W(p_2) = 0$ and to prove that $W \equiv 0$.

Suppose that W achieves its maximum at \hat{p} . Then $W'(\hat{p}) = 0$, therefore:

$$\begin{aligned} W(\hat{p}) &= \frac{1}{r} \left[\int (\hat{p}\nu_1(dh) + (1 - \hat{p})\nu_2(dh)) W \left(\frac{\hat{p}\nu_1(dh)}{\hat{p}\nu_1(dh) + (1 - \hat{p})\nu_2(dh)} \right) \right. \\ &\quad \left. + \frac{1}{2} W''(\hat{p}) \hat{p}^2 (1 - \hat{p})^2 (\tilde{\mu}_1 - \tilde{\mu}_2)^2 - (\hat{p}\bar{\nu}_1 + (1 - \hat{p})\bar{\nu}_2) W(\hat{p}) \right]. \end{aligned}$$

Moreover, since the maximum is achieved at \hat{p} , $W''(\hat{p}) \leq 0$, simple algebraic manipulations imply that:

$$\begin{aligned} (r + \hat{p}\bar{\nu}_1 + (1 - \hat{p})\bar{\nu}_2) W(\hat{p}) &= \int (\hat{p}\nu_1(dh) + (1 - \hat{p})\nu_2(dh)) W \left(\frac{\hat{p}\nu_1(dh)}{\hat{p}\nu_1(dh) + (1 - \hat{p})\nu_2(dh)} \right) \\ &\quad + \frac{1}{2} W''(\hat{p}) \hat{p}^2 (1 - \hat{p})^2 (\tilde{\mu}_1 - \tilde{\mu}_2)^2 \\ &\leq W(\hat{p}) \int (\hat{p}\nu_1(dh) + (1 - \hat{p})\nu_2(dh)) \\ &= (\hat{p}\bar{\nu}_1 + (1 - \hat{p})\bar{\nu}_2) W(\hat{p}). \end{aligned}$$

Since $r > 0$ we conclude that $W(\hat{p}) = 0$. A similar argument shows that the minimum of W in (p_1, p_2) is 0, so that $W(p) \equiv 0$ on (p_1, p_2) , as desired. \square

As mentioned before, there is an optimal cut-off strategy with corresponding payoff U . Lemmas 4.3, 4.4 and 4.5 prove that U is the unique solution of Eq. (FDE). We now prove Theorem 2.3, which provides an explicit form to the optimal strategy and to the payoff function.

Proof of Theorem 2.3.

Recall that κ^* is the optimal cut-off strategy (with cut-off point p^*). To complete the proof of the theorem, we provide an explicit expression to p^* and to U . To this end, we first prove that the right-hand derivative of U at p^* is 0.

Let $U'_R(p^*) = \lim_{\epsilon \rightarrow 0^+} \frac{U(p^* + \epsilon) - U(p^*)}{\epsilon}$ be the right derivative of U at the cut-off point p^* . Since U is convex, $U'_R(p^*)$ is well defined. Since U is non-decreasing, $U'_R(p^*) \geq 0$. We now prove that $U'_R(p^*) \leq 0$. For every $q_0 \in [0, 1]$, let $\kappa(q_0)$ be the strategy that plays as κ , assuming the prior belief is q_0 rather than p_0 . Define $M_\kappa := \int_0^\infty r e^{-rt} dY^\kappa(t)$, the discounted payoff under the strategy κ .

$$\lim_{\epsilon \rightarrow 0^+} E[M_{\kappa^*(p^* + \epsilon)} | \theta] = s \quad \forall \theta \in \{\theta_1, \theta_2\}.$$

Indeed, as explained before, the posterior belief will drop below p_0 in an infinitesimal time interval around 0. Therefore,¹⁵ as ϵ goes to 0, the probability that the DM stops “quite fast” under $\kappa^*(p^* + \epsilon)$ goes to 1.

Since κ^* is the optimal cut-off strategy, and since it is independent of the prior belief p_0 , we deduce that for every $p \in [0, 1]$, $U(p) = V_{\kappa^*(p)}(p)$. Therefore:

$$\begin{aligned} (4.12) \quad U'_R(p^*) &= \lim_{\epsilon \rightarrow 0^+} \frac{U(p^* + \epsilon) - U(p^*)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{V_{\kappa^*(p^* + \epsilon)}(p^* + \epsilon) - V_{\kappa^*(p^*)}(p^*)}{\epsilon} \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} [(p^* + \epsilon)E[M_{\kappa^*(p^* + \epsilon)} | \theta_1] + (1 - p^* - \epsilon)E[M_{\kappa^*(p^* + \epsilon)} | \theta_2]] \\ &\quad - p^*E[M_{\kappa^*(p^*)} | \theta_1] - (1 - p^*)E[M_{\kappa^*(p^*)} | \theta_2]] \\ &= \lim_{\epsilon \rightarrow 0^+} \left\{ \frac{1}{\epsilon} [p^*E[M_{\kappa^*(p^* + \epsilon)} | \theta_1] + (1 - p^*)E[M_{\kappa^*(p^* + \epsilon)} | \theta_2]] \right. \\ &\quad \left. - p^*E[M_{\kappa^*(p^*)} | \theta_1] - (1 - p^*)E[M_{\kappa^*(p^*)} | \theta_2]] \right. \\ &\quad \left. + E[M_{\kappa^*(p^* + \epsilon)} | \theta_1] - E[M_{\kappa^*(p^* + \epsilon)} | \theta_2] \right\}. \end{aligned}$$

By the optimality of $U(p)$, $U(p^*) \geq V_{\kappa^*(p^* + \epsilon)}(p^*)$, and therefore,

$$\begin{aligned} (4.13) \quad &\lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} [p^*E[M_{\kappa^*(p^* + \epsilon)} | \theta_1] + (1 - p^*)E[M_{\kappa^*(p^* + \epsilon)} | \theta_2] - p^*E[M_{\kappa^*(p^*)} | \theta_1] - (1 - p^*)E[M_{\kappa^*(p^*)} | \theta_2]] \\ &= \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} [V_{\kappa^*(p^* + \epsilon)}(p^*) - V_{\kappa^*(p^*)}(p^*)] = \lim_{\epsilon \rightarrow 0^+} \frac{1}{\epsilon} [V_{\kappa^*(p^* + \epsilon)}(p^*) - U(p^*)] \leq 0, \end{aligned}$$

and

$$(4.14) \quad \lim_{\epsilon \rightarrow 0^+} (E[M_{\kappa^*(p^* + \epsilon)} | \theta_1] - E[M_{\kappa^*(p^* + \epsilon)} | \theta_2]) = 0.$$

Substituting (4.13) and (4.14) in (4.12) we deduce that $U'_R(p^*) \leq 0$.

¹⁵By the same argument we get that $V_{\kappa'}(p)$ is continuous in p' , where κ' is a cut-off strategy with cut-off point p' .

As mentioned in Lemma 2.2, $U(p^*) = s$, and $U(1) = g_1$. By Lemma 4.5, the unique solution of Eq. (FDE) on $(p^*, 1]$ is $U(p) = pg_1 + (1-p)g_2 + C_\alpha(1-p) \left(\frac{1-p}{p}\right)^\alpha$. By imposing $U(p^*) = s$ and $U'_R(p^*) = 0$ we get $p^* = \frac{\alpha(s-g_2)}{(\alpha+1)(g_1-s)+\alpha(s-g_2)}$, and $C_\alpha = \frac{s-g_2-p^*(g_1-g_2)}{(1-p^*)(\frac{1-p^*}{p^*})^\alpha}$. Uniqueness follows by (4.10). \square

4.4. Incorrect Prior. To find the expected discounted payoff for a DM who plays according to an incorrect prior q_0 , we present here a condition which is equivalent to (2.6). Substituting $Y_B^1(t) = \mu t + \sigma Z(t)$ and $\tilde{\mu} = \frac{\mu}{\sigma}$ in (2.6), we get:

$$Z(t) + \left[\frac{2\tilde{\mu} - \tilde{\mu}_1 - \tilde{\mu}_2}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu}_1 - \tilde{\mu}_2} \right] t > -\frac{1}{\tilde{\mu}_1 - \tilde{\mu}_2} \left[\ln \left(\frac{q_0}{1-q_0} \right) - \ln \left(\frac{p'}{1-p'} \right) \right] - \frac{1}{\tilde{\mu}_1 - \tilde{\mu}_2} \sum_{t-} \ln \left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)} \right).$$

It follows that the DM selects the risky arm until the first time t is satisfied:

(4.15)

$$\begin{aligned} B^\mu(t) &:= Z(t) + \left[\frac{2\tilde{\mu} - \tilde{\mu}_1 - \tilde{\mu}_2}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu}_1 - \tilde{\mu}_2} \right] t \\ &\leq -\frac{1}{\tilde{\mu}_1 - \tilde{\mu}_2} \left[\ln \left(\frac{q_0}{1-q_0} \right) - \ln \left(\frac{p'}{1-p'} \right) \right] - \frac{1}{\tilde{\mu}_1 - \tilde{\mu}_2} \sum_{t-} \ln \left(\frac{\nu_1(dh_j)}{\nu_2(dh_j)} \right). \end{aligned}$$

Proof of Theorem 2.5. If the prior belief of the DM, q_0 , satisfies $q_0 \geq p'$ then there is a bijection relation between E and q_0 . If $q_0 \leq p'$, then the DM always chooses the safe arm, which is equivalent to $E = 0$. Therefore, we will use the notation $U(p_0, E)$ instead of $U(p_0, q_0)$ when the former is more convenient. We now prove that under Assumption 2.1, for every $p_0 \in [0, 1]$ and every $E \in [0, \infty)$, the payoff of the DM is

$$U(p_0, E) = p_0 g_1 + (1-p_0) g_2 + (s-g_1) p_0 e^{-(\tilde{\mu}_1 - \tilde{\mu}_2)(\alpha+1)E} + (s-g_2)(1-p_0) e^{-(\tilde{\mu}_1 - \tilde{\mu}_2)\alpha E}.$$

Using Eq. (4.15), we construct an integral equation, to find the utility function for a DM who has a prior belief q_0 . Let τ be the stopping time of the first jump. Let T be the first time t satisfying (4.15). The DM chooses the risky arm until the stopping time T , and then he switches to the safe arm. The calculations use dynamic programming in which the continuation payoff is determined by the time of the first jump, τ , and the value of the continuous part of the payoff at that time.

Notation and Formulas. The proof requires computations that rely on some results on Brownian motion. In this subsection we provide these results, that are derived using Borodin and Salminen (1996) p.197 - 223. Recall that, $\tilde{\mu} = \frac{\mu}{\sigma}$ is determined by θ . Notice that $B^\mu(t)$ is a standard Brownian motion with drift $F_\mu := \left[\frac{2\tilde{\mu} - \tilde{\mu}_1 - \tilde{\mu}_2}{2} - \frac{\bar{\nu}_1 - \bar{\nu}_2}{\tilde{\mu}_1 - \tilde{\mu}_2} \right] t$. Define $f_{(B^{\mu_i}(\tau), \tau) | \tau < T}^{\theta_i}(x, t) := P_\theta(B^\mu(\tau) \in dx, \tau = t | \tau < T)$. This is the probability that the first jump occurs in the interval $[t, t+dt)$, and B^μ belongs to $[x, x+dx)$, given a jump occurs before the DM switches to the safe arm.

Denote

(4.16)

$$\begin{aligned}
p_{t,h,x} &:= P(\theta_1 | \tau < T, \tau = t, B^\mu(\tau) \in dx, h) \\
&= \frac{P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h | \theta_1) P(\theta_1)}{P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h | \theta_1) P(\theta_1) + P(\tau < T, \tau = t, B^\mu(\tau) \in dx, h | \theta_2) P(\theta_2)} \\
&= \frac{p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1}{p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1 + (1 - p_0) P_{\theta_2}(\tau < T) f_{(B^{\mu_2}(\tau), \tau) | \tau < T}^{\theta_2}(x, t) \nu_2(dh) / \bar{\nu}_2}.
\end{aligned}$$

This is the posterior belief that the type is θ_1 , given that (a) the first jump that occurred in the time interval $[t, t + dt]$ has size h ; (b) it occurred before the DM switched to the safe arm; and (c) the Brownian motion with drift, $B^\mu(t)$, is in the interval $[x, x + dx]$.

The probability that the DM switches to the safe arm before the first jump appeared is

$$(4.17) \quad P_\theta(\tau > T) = P_\theta\left(\inf_{0 < s < \tau} B^\mu(s) \leq -E\right) = e^{-E(F_\mu + \sqrt{2\bar{\nu} + F_\mu^2})}.$$

The expected discounted payoff from the continuous part of the risky arm, until the switching time to the safe arm, given the DM switched before the first jump, is

(4.18)

$$\begin{aligned}
E_\theta \left[\int_0^T r e^{-rt} dY_B^1(t) \middle| \tau > T \right] &= \mu E_\theta \left[\int_0^T r e^{-rt} dt \middle| \tau > T \right] + \sigma E_\theta \left[\int_0^T r e^{-rt} dZ_t \middle| \tau > T \right] \\
&= \mu E_\theta [1 - e^{-rT} | \tau > T] = \mu(1 - E_\theta[e^{-rT} | \tau > T]).
\end{aligned}$$

The expected discounted payoff from the safe arm, after the switching time to the safe arm, given the DM switched before the first jump, is

$$(4.19) \quad E_\theta \left[\int_T^\infty r e^{-rt} s dt \middle| \tau > T \right] = s E_\theta [e^{-rT} | \tau > T].$$

The expected discounted payoff from the continuous part of the risky arm, until the first jump occurs, given the first jump occurred before the switching time, is

$$(4.20) \quad E_\theta \left[\int_0^\tau r e^{-rt} dY_B^1(t) \middle| \tau < T \right] = \mu(1 - E_\theta[e^{-r\tau} | \tau < T]).$$

The expression on the right-hand side of (4.19), can be re-written using the following list of equalities:

(4.21)

$$\begin{aligned}
E_\theta[e^{-rT}|\tau > T] &= \frac{E_\theta[e^{-rT}, \tau > T]}{P_\theta(\tau > T)} = \frac{\int e^{-rt_1} P_\theta(T \in dt_1, t_1 < \tau) dt_1}{P_\theta(\tau > T)} \\
&= \frac{\int e^{-rt_1} P_\theta(T \in dt_1) P_\theta(t_1 < \tau) dt_1}{P_\theta(\tau > T)} = \frac{\int e^{-rt_1} P_\theta(T \in dt_1) e^{-\bar{\nu}t_1} dt_1}{P_\theta(\tau > T)} \\
&= \frac{\int e^{-(r+\bar{\nu})t_1} P_\theta(T \in dt_1) dt_1}{P_\theta(\tau > T)} = \frac{E_\theta[e^{-(r+\bar{\nu})T}]}{P_\theta(\tau > T)} \\
&= \frac{e^{-E(F_\mu + \sqrt{2(r+\bar{\nu})+F_\mu^2})}}{P_\theta(\tau > T)} = \frac{P_\theta(\tau^r > T)}{P_\theta(\tau > T)},
\end{aligned}$$

where $P_\theta(\tau^r > t) = e^{-(r+\bar{\nu})t}$, and $P_\theta(\tau^r > T) = e^{-E(F_\mu + \sqrt{2(\bar{\nu}+r)+F_\mu^2})}$. Similarly, the expression on the right-hand side of (4.20) can be re-written as follows:

(4.22)

$$\begin{aligned}
E_\theta[e^{-r\tau}|\tau < T] &= \int e^{-rt_1} P_\theta(\tau \in dt_1|\tau < T) dt_1 = \int \frac{e^{-rt_1} P_\theta(\tau \in dt_1, t_1 < T)}{P_\theta(\tau < T)} dt_1 \\
&= \int \frac{e^{-rt_1} P_\theta(\tau \in dt_1) P_\theta(t_1 < T)}{P_\theta(\tau < T)} dt_1 = \int \frac{e^{-rt_1} \bar{\nu} e^{-\bar{\nu}t_1} P_\theta(t_1 < T)}{P_\theta(\tau < T)} dt_1 \\
&= \frac{\bar{\nu}}{\bar{\nu} + r} \int (\bar{\nu} + r) \frac{e^{-(r+\bar{\nu})t_1} P_\theta(t_1 < T)}{P_\theta(\tau < T)} dt_1 \\
&= \frac{\bar{\nu}}{\bar{\nu} + r} \int \frac{P_\theta(\tau^r \in dt_1) P_\theta(t_1 < T)}{P_\theta(\tau < T)} dt_1 = \frac{\bar{\nu}}{\bar{\nu} + r} \frac{P_\theta(\tau^r < T)}{P_\theta(\tau < T)}.
\end{aligned}$$

In the proof of Theorem 2.5 we use the following two identities:

(4.23)

$$\begin{aligned}
\int_0^\infty f_{(B^{\mu_i}(\tau), \tau) | \tau < T}^{\theta_i}(x, t) e^{-\gamma t} dt &= \frac{\int e^{-\gamma t} P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx, \tau \in dt \right)}{P_\theta(\tau < T)} \\
&= \frac{\int e^{-\gamma t} P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx \right) P_\theta(\tau \in dt) dt}{P_\theta(\tau < T)} \\
&= \frac{\int e^{-\gamma t} P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx \right) \bar{\nu} e^{-\bar{\nu} t} dt}{P_\theta(\tau < T)} \\
&= \frac{\bar{\nu}}{\bar{\nu} + \gamma} \frac{\int (\bar{\nu} + \gamma) e^{-(\gamma + \bar{\nu})t} P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx \right) dt}{P_\theta(\tau < T)} \\
&= \frac{\bar{\nu}}{\bar{\nu} + \gamma} \frac{\int P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx \right) P_\theta(\tau^\gamma \in dt) dt}{P_\theta(\tau < T)} \\
&= \frac{\bar{\nu}}{\bar{\nu} + \gamma} \frac{\int P_\theta \left(\inf_{0 \leq s \leq t} B^\mu(s) \geq -E, B^\mu(t) \in dx, \tau^\gamma \in dt \right) dt}{P_\theta(\tau < T)} \\
&= \frac{\bar{\nu}}{\bar{\nu} + \gamma} \frac{P_\theta(B^\mu(\tau^\gamma) \in dx, \tau^\gamma < T)}{P_\theta(\tau < T)}.
\end{aligned}$$

(4.24)

$$\begin{aligned}
& \int_{-E}^{\infty} \int_0^{\infty} f_{(B^{\mu_i}(\tau), \tau)}^{\theta_i}(x, t) e^{-\gamma t} e^{-\delta x} dt dx = \frac{\bar{\nu}}{\bar{\nu} + \gamma} \frac{\int_{-E}^{\infty} P_{\theta}(B^{\mu}(\tau^{\gamma}) \in dx, \tau^{\gamma} < T) e^{-\delta x} dx}{P_{\theta}(\tau < T)} \\
&= \frac{\bar{\nu}}{(\bar{\nu} + \gamma) P_{\theta}(\tau < T)} \int_{-E}^{\infty} P_{\theta} \left(\inf_{0 < s < \tau^{\gamma}} B^{\mu}(s) > -E, B^{\mu}(\tau^{\gamma}) \in dx \right) e^{-\delta x} dx \\
&= \frac{\bar{\nu}}{(\bar{\nu} + \gamma) P_{\theta}(\tau < T)} \int_{-E}^{\infty} \left[P_{\theta}(B^{\mu}(\tau^{\gamma}) \in dx) e^{-\delta x} - P_{\theta} \left(\inf_{0 < s < \tau^{\gamma}} B^{\mu}(s) \leq -E, B^{\mu}(\tau^{\gamma}) \in dx \right) e^{-\delta x} \right] dx \\
&= \frac{\bar{\nu} + \gamma}{\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2}} \frac{\bar{\nu}}{(\bar{\nu} + \gamma) P_{\theta}(\tau < T)} \cdot \int_{-E}^{\infty} \left[e^{(F_{\mu} - \delta)x - |x| \sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2}} - e^{-x(\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} + \delta - F_{\mu})} e^{-2E \sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2}} \right] dx \\
&= \frac{\bar{\nu}}{\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} P_{\theta}(\tau < T)} \cdot \left[\int_0^{\infty} e^{-x(\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} + \delta - F_{\mu})} + \int_{-E}^0 e^{x(\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} - \delta + F_{\mu})} \right. \\
&\quad \left. - e^{-2E \sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2}} \int_{-E}^{\infty} e^{-x(\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} + \delta - F_{\mu})} \right] \\
&= \frac{\bar{\nu}(1 - e^{-E(\sqrt{2(\bar{\nu} + \gamma) + F_{\mu}^2} + F_{\mu} - \delta)})}{(\bar{\nu} + \gamma + \delta F_{\mu} - \delta^2/2) P_{\theta}(\tau < T)}.
\end{aligned}$$

Where the first equality follows by (4.23).

Constructing the integral equation. The DM chooses the risky arm, until the minimum between the stopping time of the first jump τ and the stopping time T . We distinguish between two cases. In case the DM stops before the time of the first jump τ , we calculate the expected discounted payoff from the risky arm until the stopping time T , and the expected discounted payoff from the safe arm after the stopping time T . In case the first jump occurs before the stopping time T , we calculate the expected discounted payoff received from the continuous part $Y_B^1(t)$ until time τ . We add the expected discounted payoff from the first jump, and the expected discounted continuation payoff, updating both the posterior $p_{t,h,x}$, and the intercept $E + G_h + x$, according to the time of the first jump, the first jump's size, and the value of the continuous part of the payoff.

The notation used are as follows: $P_{\theta}(\tau > T)$ is the probability that the DM switches to the safe arm before a jump occurs. If $\tau > T$ then the expected payoff from the risky arm is $E_{\theta}[\int_0^T re^{-rt} dY_B^1(t) | \tau > T]$, and the expected payoff from the safe arm is $E_{\theta}[\int_T^{\infty} re^{-rt} s dt | \tau > T]$. $P_{\theta}(\tau < T)$ is the probability a jump occurs before the DM switches to the safe arm. If $\tau < T$ then the expected payoff until the first jump is $E_{\theta}[\int_0^{\tau} re^{-rt} dY_B^1(t) | \tau < T]$. $f_{(B^{\mu_i}(\tau), \tau)}^{\theta_i}(x, t)$ is the probability that the first jump occurs in the interval $[t, t + dt)$, and $B^{\mu}(\tau)$ belongs to $[x, x + dx)$, given that a jump occurs before the DM switches to the safe arm. $re^{-rt} \frac{1}{\bar{\nu}} \int h \nu(dh)$ is the expected discounted payoff from the first jump, and $\frac{1}{\bar{\nu}} \int \nu e^{-rt} U(p_{t,h,x}, E + G_h + x)$ is the expected discounted continuation payoff, updating both the posterior $p_{t,h,x}$, and the intercept $E + G_h + x$ at time t . With this notation, the expected payoff is

as follows:

(4.25)

$$\begin{aligned}
U(p_0, E) &= p_0 P_{\theta_1}(\tau < T) \left[E_{\theta_1} \left[\int_0^\tau r e^{-rt} dY_B^1(t) \mid \tau < T \right] \right. \\
&\quad \left. + \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_1}(\tau), \tau)}^{\theta_1} |_{\tau < T}(x, t) \left(r e^{-rt} \frac{1}{\bar{\nu}_1} \int \nu_1(dh) h + \frac{1}{\bar{\nu}_1} \int \nu_1(dh) e^{-rt} U(p_{t,h,x}, E + G_h + x) \right) dt dx \right] \\
&\quad + p_0 P_{\theta_1}(\tau > T) \left[E_{\theta_1} \left[\int_0^T r e^{-rt} dY_B^1(t) \mid \tau > T \right] + E_{\theta_1} \left[\int_T^\infty r e^{-rt} s dt \mid \tau > T \right] \right] \\
&\quad + (1 - p_0) P_{\theta_2}(\tau < T) \left[E_{\theta_2} \left[\int_0^\tau r e^{-rt} dY_B^1(t) \mid \tau < T \right] \right. \\
&\quad \left. + \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_2}(\tau), \tau)}^{\theta_2} |_{\tau < T}(x, t) \left(r e^{-rt} \frac{1}{\bar{\nu}_2} \int \nu_2(dh) h + \frac{1}{\bar{\nu}_2} \int \nu_2(dh) e^{-rt} U(p_{t,h,x}, E + G_h + x) \right) dt dx \right] \\
&\quad + (1 - p_0) P_{\theta_2}(\tau > T) \left[E_{\theta_2} \left[\int_0^T r e^{-rt} dY_B^1(t) \mid \tau > T \right] + E_{\theta_2} \left[\int_T^\infty r e^{-rt} s dt \mid \tau > T \right] \right].
\end{aligned}$$

By Eqs. (4.18), (4.19) and (4.20), this expression is equal to

$$\begin{aligned}
&\mu_1 p_0 (1 - P_{\theta_1}(\tau > T)) (1 - E_{\theta_1}[e^{-r\tau} \mid \tau < T]) \\
&\quad + p_0 P_{\theta_1}(\tau < T) H_1 \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_1}(\tau), \tau)}^{\theta_1} |_{\tau < T}(x, t) r e^{-rt} dt dx \\
&\quad + p_0 P_{\theta_1}(\tau < T) \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_1}(\tau), \tau)}^{\theta_1} |_{\tau < T}(x, t) \frac{1}{\bar{\nu}_1} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \nu_1(dh) dt dx \\
&\quad + \mu_1 p_0 P_{\theta_1}(\tau > T) (1 - E_{\theta_1}[e^{-rT} \mid \tau > T]) + s p_0 P_{\theta_1}(\tau > T) E_{\theta_1}[e^{-rT} \mid \tau > T] \\
&\quad + \mu_2 (1 - p_0) P_{\theta_2}(\tau > T) (1 - E_{\theta_2}[e^{-r\tau} \mid \tau < T]) \\
&\quad + (1 - p_0) P_{\theta_2}(\tau < T) H_2 \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_2}(\tau), \tau)}^{\theta_2} |_{\tau < T}(x, t) r e^{-rt} dt dx \\
&\quad + (1 - p_0) P_{\theta_2}(\tau < T) \int_{-E}^\infty \int_0^\infty f_{(B^{\mu_2}(\tau), \tau)}^{\theta_2} |_{\tau < T}(x, t) \frac{1}{\bar{\nu}_2} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \nu_2(dh) dt dx \\
&\quad + \mu_2 (1 - p_0) P_{\theta_2}(\tau > T) (1 - E_{\theta_2}[e^{-rT} \mid \tau > T]) + s (1 - p_0) P_{\theta_2}(\tau > T) E_{\theta_2}[e^{-rT} \mid \tau > T].
\end{aligned}$$

By Eqs. (4.21), (4.22) and (4.23), this expression is equal to

$$\begin{aligned}
&\mu_1 p_0 P_{\theta_1}(\tau > T) \left(1 - \frac{P_{\theta_1}(\tau^r > T)}{P_{\theta_1}(\tau > T)} \right) + s p_0 P_{\theta_1}(\tau > T) \frac{P_{\theta_1}(\tau^r > T)}{P_{\theta_1}(\tau > T)} \\
&\quad + p_0 (1 - P_{\theta_1}(\tau > T)) H_1 r \cdot \frac{\bar{\nu}_1}{\bar{\nu}_1 + r} \cdot \frac{1 - P_{\theta_1}(\tau^r > T)}{1 - P_{\theta_1}(\tau > T)} \\
&\quad + \mu_1 p_0 (1 - P_{\theta_1}(\tau > T)) \left(1 - \frac{\bar{\nu}_1}{\bar{\nu}_1 + r} \cdot \frac{1 - P_{\theta_1}(\tau^r > T)}{1 - P_{\theta_1}(\tau > T)} \right) \\
&\quad + \mu_2 (1 - p_0) P_{\theta_2}(\tau > T) \left(1 - \frac{P_{\theta_2}(\tau^r > T)}{P_{\theta_2}(\tau > T)} \right) + s (1 - p_0) P_{\theta_2}(\tau > T) \frac{P_{\theta_2}(\tau^r > T)}{P_{\theta_2}(\tau > T)} \\
&\quad + (1 - p_0) (1 - P_{\theta_2}(\tau > T)) H_2 r \cdot \frac{\bar{\nu}_2}{\bar{\nu}_2 + r} \cdot \frac{1 - P_{\theta_2}(\tau^r > T)}{1 - P_{\theta_2}(\tau > T)}
\end{aligned}$$

$$\begin{aligned}
& + \mu_2(1-p_0)(1-P_{\theta_2}(\tau > T)) \left(1 - \frac{\bar{\nu}_1}{\bar{\nu}_1+r} \cdot \frac{1-P_{\theta_2}(\tau^r > T)}{1-P_{\theta_2}(\tau > T)}\right) \\
& + \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \cdot \\
& \cdot \left[p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau)|\tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1 + (1-p_0) P_{\theta_2}(\tau < T) f_{(B^{\mu_2}(\tau), \tau)|\tau < T}^{\theta_2}(x, t) \nu_2(dh) / \bar{\nu}_2 \right] dt dx \\
& = p_0 \left(\frac{\mu_1 r}{\bar{\nu}_1+r} + \frac{\bar{\nu}_1 H_1 r}{\bar{\nu}_1+r} \right) + (1-p_0) \left(\frac{\mu_2 r}{\bar{\nu}_2+r} + \frac{\bar{\nu}_2 H_2 r}{\bar{\nu}_2+r} \right) \\
& + p_0 P_{\theta_1}(\tau^r > T) \left(s - \frac{\mu_1 r}{\bar{\nu}_1+r} - \frac{\bar{\nu}_1 H_1 r}{\bar{\nu}_1+r} \right) + (1-p_0) P_{\theta_2}(\tau^r > T) \left(s - \frac{\mu_2 r}{\bar{\nu}_2+r} - \frac{\bar{\nu}_2 H_2 r}{\bar{\nu}_2+r} \right) \\
& + \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \cdot \\
& \cdot \left[p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau)|\tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1 + (1-p_0) P_{\theta_2}(\tau < T) f_{(B^{\mu_2}(\tau), \tau)|\tau < T}^{\theta_2}(x, t) \nu_2(dh) / \bar{\nu}_2 \right] dt dx \\
& = p_0 \left(\frac{\mu_1 r}{\bar{\nu}_1+r} + \frac{\bar{\nu}_1 H_1 r}{\bar{\nu}_1+r} \right) + (1-p_0) \left(\frac{\mu_2 r}{\bar{\nu}_2+r} + \frac{\bar{\nu}_2 H_2 r}{\bar{\nu}_2+r} \right) \\
& + p_0 e^{-E(F_{\mu_1} + \sqrt{2(\bar{\nu}_1+r)+F_{\mu_1}^2})} \left(s - \frac{\mu_1 r}{\bar{\nu}_1+r} - \frac{\bar{\nu}_1 H_1 r}{\bar{\nu}_1+r} \right) \\
& + (1-p_0) e^{-E(F_{\mu_2} + \sqrt{2(\bar{\nu}_2+r)+F_{\mu_2}^2})} \left(s - \frac{\mu_2 r}{\bar{\nu}_2+r} - \frac{\bar{\nu}_2 H_2 r}{\bar{\nu}_2+r} \right) \\
& + \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) \cdot \\
& \cdot \left[p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau)|\tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1 + (1-p_0) P_{\theta_2}(\tau < T) f_{(B^{\mu_2}(\tau), \tau)|\tau < T}^{\theta_2}(x, t) \nu_2(dh) / \bar{\nu}_2 \right] dt dx.
\end{aligned}$$

Simplifying the last expression, the integral equation is

$$\begin{aligned}
(IE) \quad U(p, E) &= Ap + B(1-p) + Cpe^{-m_1 E} + D(1-p)e^{-m_2 E} \\
&+ \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} U(p_{t,h,x}, E + G_h + x) g(x, t, h) dt dx,
\end{aligned}$$

where A, B, C , and D are constants, and

$$\begin{aligned}
(4.26) \quad g(x, t, h) &= \left[p_0 P_{\theta_1}(\tau < T) f_{(B^{\mu_1}(\tau), \tau)|\tau < T}^{\theta_1}(x, t) \nu_1(dh) / \bar{\nu}_1 \right. \\
&\quad \left. + (1-p_0) P_{\theta_2}(\tau < T) f_{(B^{\mu_2}(\tau), \tau)|\tau < T}^{\theta_2}(x, t) \nu_2(dh) / \bar{\nu}_2 \right].
\end{aligned}$$

We show now that (IE) admits a unique solution U in the region $[0, 1] \times [0, \infty)$.

Boundary values

First, we find the values of $U(p, E)$ on the boundary of the region $[0, 1] \times [0, \infty]$. Note that $U(p, 0) = s$, and $U(p, \infty) = pg_1 + (1-p)g_2$ for every $p \in [0, 1]$.

We now argue that there is a unique solution for (IE) when $p = 0$. $U(0, E)$ is a function of E with two boundary conditions, at $E = 0$ and at $E = \infty$. Suppose that $U(0, E), V(0, E)$ solve (IE). Then $W(0, E) := U(0, E) - V(0, E)$ satisfies $W(0, E) = \int_{-E}^{\infty} \int_0^{\infty} \int e^{-rt} W(0, E + G_h + x) g(x, t, h) dt dx$, and $W(0, 0) = W(0, \infty) = 0$. Let \hat{E} be a critical point, where W achieves its maximum. Assume to the contrary that

$W(0, \hat{E}) > 0$. By (4.23) it follows that $\int_{-\hat{E}}^{\infty} \int_0^{\infty} \int e^{-rt} g(x, t, h) < 1$. Therefore,

$$\begin{aligned} W(0, \hat{E}) &= \int_{-\hat{E}}^{\infty} \int_0^{\infty} \int e^{-rt} W(0, \hat{E} + G_h + x) g(x, t, h) \\ &\leq W(0, \hat{E}) \int_{-\hat{E}}^{\infty} \int_0^{\infty} \int e^{-rt} g(x, t, h) < W(0, \hat{E}), \end{aligned}$$

which implies that $W(0, E) \leq 0$. Similarly, one can obtain that $W(0, E) \geq 0$, so that $W(0, E) \equiv 0$, and the solution is unique.

Similar arguments show that (IE) admits a unique solution on $[0, \infty)$ when $p = 1$.

Since $U(p, E)$ is uniquely determined on the boundary of the region $[0, 1] \times [0, \infty]$, similar arguments show the uniqueness of the solution $[0, 1] \times [0, \infty)$. Using Eqs. (4.16) - (4.24) one can verify that the solution for (IE) is

$$U(p_0, E) = p_0 g_1 + (1 - p_0) g_2 + (s - g_1) p_0 e^{-(\hat{\mu}_1 - \hat{\mu}_2)(\alpha + 1)E} + (s - g_2)(1 - p_0) e^{-(\hat{\mu}_1 - \hat{\mu}_2)\alpha E}.$$

By substituting $E := \frac{\sigma}{\mu_1 - \mu_2} \times \left[\ln \left(\frac{q_0}{1 - q_0} \right) - \ln \left(\frac{p'}{1 - p'} \right) \right]$, we get Eq. (2.7), as desired. \square

4.5. Information Pricing.

Lemma 4.6. *Let $f_{a,b}(\eta)$ be the function defined in (2.8). The equation $f_{a,b}(\eta) = 0$ admits a unique solution in the interval $(0, \infty)$.*

The proof is similar to the proof of Lemma 4.4, and therefore omitted. We turn to the proof of Theorem 2.6, which is analogous to the proof of Theorems 2.3 and 2.5.

Proof of Theorem 2.6. Suppose that until time t , the DM chose the risky arm, and observed the jumps h_1^a, \dots, h_n^a (resp. h_1^b, \dots, h_m^b) from the process $(X^a(t))$ (resp. $(X^b(t))$). Let $Y_B^a(t)$ (resp. $Y_B^b(t)$) be the Brownian motion with drift component of $(X^a(t))$ (resp. $(X^b(t))$) at time t . Note that $Y_B^j(t) \sim N(\mu^j t, (\sigma^j)^2 t)$, $j \in \{a, b\}$. Let $q_t := P_t(\theta_1 | h_1, \dots, h_n; h_1^b, \dots, h_m^b; Y_B^a(t); Y_B^b(t); q_0)$ be the posterior belief of the DM. The odd ratio of the posterior belief is

$$\begin{aligned} \frac{q_t}{1 - q_t} &= \frac{q_0 \frac{1}{\sqrt{2\pi t \sigma^a}} e^{-\frac{(Y_B^a(t) - \mu_1^a t)^2}{2(\sigma^a)^2 t}} \frac{1}{\sqrt{2\pi t \sigma^b}} e^{-\frac{(Y_B^b(t) - \mu_1^b t)^2}{2(\sigma^b)^2 t}} e^{-\bar{\nu}_1^a t} \prod_{t-} \nu_1^a(dh_j^a) e^{-\bar{\nu}_1^b t} \prod_{t-} \nu_1^b(dh_j^b)}{(1 - q_0) \frac{1}{\sqrt{2\pi t (\sigma^a)^2}} e^{-\frac{(Y_B^a(t) - \mu_2^a t)^2}{2(\sigma^a)^2 t}} \frac{1}{\sqrt{2\pi t (\sigma^b)^2}} e^{-\frac{(Y_B^b(t) - \mu_2^b t)^2}{2(\sigma^b)^2 t}} e^{-\bar{\nu}_2^a t} \prod_{t-} \nu_2^a(dh_j^a) e^{-\bar{\nu}_2^b t} \prod_{t-} \nu_2^b(dh_j^b)} \\ &= \frac{q_0 e^{\mu_1^a Y_B^a(t)/(\sigma^a)^2 - (\mu_1^a)^2 t/2(\sigma^a)^2} e^{\mu_1^b Y_B^b(t)/(\sigma^b)^2 - (\mu_1^b)^2 t/2(\sigma^b)^2} e^{-\bar{\nu}_1^a t} \prod_{t-} \nu_1^a(dh_j^a) e^{-\bar{\nu}_1^b t} \prod_{t-} \nu_1^b(dh_j^b)}{(1 - q_0) e^{\mu_2^a Y_B^a(t)/(\sigma^a)^2 - (\mu_2^a)^2 t/2(\sigma^a)^2} e^{\mu_2^b Y_B^b(t)/(\sigma^b)^2 - (\mu_2^b)^2 t/2(\sigma^b)^2} e^{-\bar{\nu}_2^a t} \prod_{t-} \nu_2^a(dh_j^a) e^{-\bar{\nu}_2^b t} \prod_{t-} \nu_2^b(dh_j^b)}. \end{aligned}$$

Indeed, $\frac{1}{\sqrt{2\pi t \sigma^a}} e^{-\frac{(Y_B^a(t) - \mu_1^a t)^2}{2(\sigma^a)^2 t}}$ (resp. $\frac{1}{\sqrt{2\pi t \sigma^b}} e^{-\frac{(Y_B^b(t) - \mu_1^b t)^2}{2(\sigma^b)^2 t}}$) is the probability of observing $Y_B^a(t)$ (resp. $Y_B^b(t)$), given the type θ_i , and $e^{-\bar{\nu}_i^a t} \frac{(\bar{\nu}_i^a t)^n}{n!} \prod_{t-} \frac{\nu_i^a(dh_j^a)}{\bar{\nu}_i^a}$ (resp. $e^{-\bar{\nu}_i^b t} \frac{(\bar{\nu}_i^b t)^m}{m!} \prod_{t-} \frac{\nu_i^b(dh_j^b)}{\bar{\nu}_i^b}$) is the probability of receiving the n (resp. m) jumps that occurred until time t from $(X_i^a(t))$ (resp. $(X_i^b(t))$), given the type θ_i . The first equality in (4.27) is the Bayesian belief updating, using the independence of the Lévy processes $X_i^a(t)$ and $X_i^b(t)$, and the independence of the components in the Lévy-Ito decomposition, given the type of the risky arm, and the second equality

is obtained by eliminating common components.

Suppose the DM follows a cut-off strategy κ' with cut-off point p' . If $q_0 \leq p'$, the DM will always choose the safe arm. If $q_0 > p'$ the DM will initially choose the risky arm. Then the DM chooses the risky arm as long as $q_t > p'$, which, by Eq. (2.4), is equivalent to

$$(4.28) \quad \frac{q_0}{1 - q_0} \times \frac{e^{\mu_1^a Y_B^a(t)/(\sigma^a)^2 - (\mu_1^a)^2 t/2(\sigma^a)^2 - \bar{\nu}_1^a t}}{e^{\mu_2^a Y_B^a(t)/(\sigma^a)^2 - (\mu_2^a)^2 t/2(\sigma^a)^2 - \bar{\nu}_2^a t}} \times \frac{e^{\mu_1^b Y_B^b(t)/(\sigma^b)^2 - (\mu_1^b)^2 t/2(\sigma^b)^2 - \bar{\nu}_1^b t}}{e^{\mu_2^b Y_B^b(t)/(\sigma^b)^2 - (\mu_2^b)^2 t/2(\sigma^b)^2 - \bar{\nu}_2^b t}} \\ \times \prod_{t-} \frac{\nu_1^a(dh^a)}{\nu_2^a(dh^a)} \times \prod_{t-} \frac{\nu_1^b(dh^b)}{\nu_2^b(dh^b)} > \frac{p'}{1 - p'}.$$

By taking the natural logarithm, and rearranging the resulting terms, we obtain that this inequality is equivalent to

$$(4.29) \quad \left(\frac{\mu_1^a - \mu_2^a}{\sigma^a} \right) Y_B^a(t) + \left(\frac{\mu_1^b - \mu_2^b}{\sigma^b} \right) Y_B^b(t) > \\ \left(\frac{(\mu_1^a)^2 - (\mu_2^a)^2}{2(\sigma^a)^2} + \frac{(\mu_1^b)^2 - (\mu_2^b)^2}{2(\sigma^b)^2} + \bar{\nu}_1^a + \bar{\nu}_1^b - \bar{\nu}_2^a - \bar{\nu}_2^b \right) t \\ - \left(\ln \left(\frac{q_0}{1 - q_0} \right) - \ln \left(\frac{p'}{1 - p'} \right) \right) - \sum_{t-} \ln \left(\frac{\nu_1^a(dh_j^a)}{\nu_2^a(dh_j^a)} \right) - \sum_{t-} \ln \left(\frac{\nu_1^b(dh_j^b)}{\nu_2^b(dh_j^b)} \right).$$

Since $Y_B^j(t) = \mu^a + \sigma^a Z^a(t)$, $k \in \{a, b\}$, it follows that Eq. (4.29) is equivalent to

$$(4.30) \quad B^{\tilde{\mu}^a, \tilde{\mu}^b}(t) := \frac{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)Z^a(t) + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)Z^b(t)}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}} + \left[\frac{-(\bar{\nu}_1^a + \bar{\nu}_1^b - \bar{\nu}_2^a - \bar{\nu}_2^b)t}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}} \right. \\ \left. + \frac{2\tilde{\mu}^a(\tilde{\mu}_1^a - \tilde{\mu}_2^a) - ((\tilde{\mu}_1^a)^2 - (\tilde{\mu}_2^a)^2) + 2\tilde{\mu}^b(\tilde{\mu}_1^b - \tilde{\mu}_2^b) - ((\tilde{\mu}_1^b)^2 - (\tilde{\mu}_2^b)^2)}{2\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}} \right] t \\ - \left(\ln \left(\frac{q_0}{1 - q_0} \right) - \ln \left(\frac{p'}{1 - p'} \right) \right) - \sum_{t-} \ln \left(\frac{\nu_1^a(dh_j^a)}{\nu_2^a(dh_j^a)} \right) - \sum_{t-} \ln \left(\frac{\nu_1^b(dh_j^b)}{\nu_2^b(dh_j^b)} \right) \\ > \frac{-(\bar{\nu}_1^a + \bar{\nu}_1^b - \bar{\nu}_2^a - \bar{\nu}_2^b)t}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}},$$

where $\tilde{\mu}_i^j = \frac{\mu_i^j}{\sigma^j}$, for $j \in \{a, b, c\}$, and $i \in \{1, 2\}$. Notice that $B^{\tilde{\mu}^a, \tilde{\mu}^b}(t)$ is a standard Brownian motion with drift:

$$F^{\tilde{\mu}^a, \tilde{\mu}^b} := \left[\frac{-(\bar{\nu}_1^a + \bar{\nu}_1^b - \bar{\nu}_2^a - \bar{\nu}_2^b)t}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}} \right. \\ \left. + \frac{2\tilde{\mu}^a(\tilde{\mu}_1^a - \tilde{\mu}_2^a) - ((\tilde{\mu}_1^a)^2 - (\tilde{\mu}_2^a)^2) + 2\tilde{\mu}^b(\tilde{\mu}_1^b - \tilde{\mu}_2^b) - ((\tilde{\mu}_1^b)^2 - (\tilde{\mu}_2^b)^2)}{2\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}} \right].$$

We construct an integral equation similar to the one in the proof of Theorem 2.5.

Let τ be the stopping time of the first jump from the process $(X^a(t) + X^b(t))$. Let the stopping time T be the first time t that satisfies

$$(4.31) \quad B^{\tilde{\mu}^a, \tilde{\mu}^b}(t) \leq \frac{-\left(\ln\left(\frac{q_0}{1-q_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right) - \sum_{t-} \ln\left(\frac{\nu_1^a(dh_j^a)}{\nu_2^a(dh_j^a)}\right) - \sum_{t-} \ln\left(\frac{\nu_1^b(dh_j^b)}{\nu_2^b(dh_j^b)}\right)}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}}.$$

Denote $G_h^a := \frac{\ln\left(\frac{\nu_1^a(dh_j^a)}{\nu_2^a(dh_j^a)}\right)}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}}$ (resp. $G_h^b := \frac{\ln\left(\frac{\nu_1^b(dh_j^b)}{\nu_2^b(dh_j^b)}\right)}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}}$) the contribution of a jump of size h^a (resp. h^b) received from the process $X^a(t)$ (resp. $X^b(t)$), and $E := \frac{-\left(\ln\left(\frac{q_0}{1-q_0}\right) - \ln\left(\frac{p'}{1-p'}\right)\right)}{\sqrt{(\tilde{\mu}_1^a - \tilde{\mu}_2^a)^2 + (\tilde{\mu}_1^b - \tilde{\mu}_2^b)^2}}$, the intercept of the right-hand side of Eq. (4.30) at $t = 0$.

Constructing the integral equation. The DM chooses the risky arm, until the minimum between the stopping time of the first jump τ and the stopping time T . We distinguish between two cases. In case the DM stops before the time of the first jump τ , we calculate the expected discounted payoff from the process $(X^a(t) + X^c(t))$ until the stopping time T , and the expected discounted payoff from the safe arm after the stopping time T . In case the first jump occurs before the stopping time T , we calculate the expected discounted payoff received from the process $(X^a(t) + X^c(t))$ until time τ . If the first jump was received from the process $(X^a(t))$, we add the expected discounted payoff from the first jump, and the expected discounted continuation payoff, updating both the posterior $p_{t,h^a,x}^a$, and the intercept $E + G_h^a + x$, according to the time of the first jump, the first jump's size, and the value of the continuous part of the process $(X^a(t) + X^b(t))$; while if the first jump was received from the process $(X^b(t))$, we add only the expected discounted continuation payoff, updating both the posterior $p_{t,h^b,x}^b$, and the intercept $E + G_h^b + x$, according to the time of the first jump, the first jump's size, and the value of the continuous part of the process $(X^a(t) + X^b(t))$.

The posterior $p_{t,h^j,x}^j$ is updated as follows:

$$\begin{aligned} p_{t,h^j,x}^j &:= P(\theta_1 | \tau < T, \tau = t, B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau) \in dx, h^j) \\ &= \frac{P(\tau < T, \tau = t, B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau) \in dx, h^j | \theta_1) P(\theta_1)}{P(\tau < T, \tau = t, B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau) \in dx, h^j | \theta_1) P(\theta_1) + P(\tau < T, \tau = t, B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau) \in dx, h^j | \theta_2) P(\theta_2)} \\ &= \frac{q_0 P_{\theta_1}(\tau < T) f_{(B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) \frac{\nu_1^j(dh^j)}{\bar{\nu}_1^j}}{q_0 P_{\theta_1}(\tau < T) f_{(B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) \frac{\nu_1^j(dh^j)}{\bar{\nu}_1^j} + (1 - q_0) P_{\theta_2}(\tau < T) f_{(B^{\tilde{\mu}^a, \tilde{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_2}(x, t) \frac{\nu_2^j(dh^j)}{\bar{\nu}_2^j}}. \end{aligned}$$

The integral equation is as follows:

(4.32)

$$\begin{aligned}
U(p_0, E) &= p_0 P_{\theta_1}(\tau < T) \left[E_{\theta_1} \left[\int_0^\tau r e^{-rt} dY_B^a(t) \middle| \tau < T \right] + E_{\theta_1} \left[\int_0^\tau r e^{-rt} dY^c(t) \middle| \tau < T \right] \right. \\
&\quad + \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) r e^{-rt} \frac{\bar{\nu}_1^a H_1^a}{\bar{\nu}_1^a + \bar{\nu}_1^b} dt dx \\
&\quad + \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) e^{-rt} \int_h U(p_{t, h^a, x}^a, E + G_h^a + x) \left(\frac{\nu_1^a(dh^a)}{\bar{\nu}_1^a + \bar{\nu}_1^b} \right) dt dx \\
&\quad + \left. \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_1}(x, t) e^{-rt} \int_h U(p_{t, h^b, x}^b, E + G_h^b + x) \left(\frac{\nu_1^b(dh^b)}{\bar{\nu}_1^a + \bar{\nu}_1^b} \right) dt dx \right] \\
&\quad + p_0 P_{\theta_1}(\tau > T) \left[E_{\theta_1} \left[\int_0^T r e^{-rt} dY_B^a(t) \middle| \tau > T \right] + E_{\theta_1} \left[\int_0^T r e^{-rt} dY^c(t) \middle| \tau > T \right] \right. \\
&\quad \quad \left. + E_{\theta_1} \left[\int_T^\infty r e^{-rt} s dt \middle| \tau > T \right] \right] \\
&\quad + (1 - p_0) P_{\theta_2}(\tau < T) \left[E_{\theta_2} \left[\int_0^\tau r e^{-rt} dY_B^a(t) \middle| \tau < T \right] + E_{\theta_2} \left[\int_0^\tau r e^{-rt} dY^c(t) \middle| \tau < T \right] \right. \\
&\quad + \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_2}(x, t) r e^{-rt} \frac{\bar{\nu}_2^a H_2^a}{\bar{\nu}_2^a + \bar{\nu}_2^b} dt dx \\
&\quad + \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_2}(x, t) e^{-rt} \int_h U(p_{t, h^a, x}^a, E + G_h^a + x) \left(\frac{\nu_2^a(dh^a)}{\bar{\nu}_2^a + \bar{\nu}_2^b} \right) dt dx \\
&\quad + \left. \int_{-E}^\infty \int_0^\infty f_{(B^{\bar{\mu}^a, \bar{\mu}^b}(\tau), \tau) | \tau < T}^{\theta_2}(x, t) e^{-rt} \int_h U(p_{t, h^b, x}^b, E + G_h^b + x) \left(\frac{\nu_2^b(dh^b)}{\bar{\nu}_2^a + \bar{\nu}_2^b} \right) dt dx \right] \\
&\quad + (1 - p_0) P_{\theta_2}(\tau > T) \left[E_{\theta_2} \left[\int_0^T r e^{-rt} dY_B^a(t) \middle| \tau > T \right] + E_{\theta_2} \left[\int_0^T r e^{-rt} dY^c(t) \middle| \tau > T \right] \right. \\
&\quad \quad \left. + E_{\theta_2} \left[\int_T^\infty r e^{-rt} s dt \middle| \tau > T \right] \right].
\end{aligned}$$

Similarly to Eqs. (4.19) and (4.20), using the stochastic integral properties for Lévy processes, it follows that for $i \in \{1, 2\}$

$$E_{\theta_i} \left[\int_0^T r e^{-rt} dY^c(t) \middle| \tau > T \right] = g_i^c (1 - E_\theta[e^{-rT} | \tau > T]),$$

and

$$E_{\theta_i} \left[\int_0^\tau r e^{-rt} dY^c(t) \middle| \tau < T \right] = g_i^c (1 - E_\theta[e^{-r\tau} | \tau < T]).$$

By these equations, and Eqs. (4.17)- (4.24), one can verify that for every $q_0 > p'$, the unique solution of Eq. (4.32) is

$$\begin{aligned}
U(p_0, E) &= (s - g_1^a - g_1^c) p_0 e^{-(\beta+1)\sqrt{(\bar{\mu}_1^a - \bar{\mu}_2^a)^2 + (\bar{\mu}_1^b - \bar{\mu}_2^b)^2} E} \\
&\quad + (s - g_2^a - g_2^c)(1 - p_0) e^{-\beta\sqrt{(\bar{\mu}_1^a - \bar{\mu}_2^a)^2 + (\bar{\mu}_1^b - \bar{\mu}_2^b)^2} E} \\
&\quad + p_0(g_1^a + g_1^c) + (1 - p_0)(g_2^a + g_2^c).
\end{aligned}$$

Therefore, for every $q_0 > p'$, the expected discounted payoff is

$$\begin{aligned} V_{p'}(p_0, q_0) &= (s - g_1^a - g_1^c) p_0 \left(\frac{1 - q_0}{q_0} \right)^{\beta+1} \left(\frac{p'}{1 - p'} \right)^{\beta+1} \\ &\quad + (s - g_2^a - g_2^c) (1 - p_0) \left(\frac{1 - q_0}{q_0} \right)^{\beta} \left(\frac{p'}{1 - p'} \right)^{\beta} \\ &\quad + p_0 (g_1^a + g_1^c) + (1 - p_0) (g_2^a + g_2^c), \end{aligned}$$

as desired.

Now, if $q_0 = p_0$, using the same method to that used to prove Theorem 2.3, one can show that the optimal cut-off strategy is

$$p^* = \frac{\beta(s - g_2^a - g_2^c)}{(\beta + 1)(g_1^a + g_1^c - s) + \beta(s - g_2^a - g_2^c)}.$$

□

Proof of Corollary 2.7. Let β_a be the unique solution of

$$f_a(\eta) := \int \nu_2^a(dh^a) \left(\frac{\nu_2^a(dh^a)}{\nu_1^a(dh^a)} \right)^{\eta} + \eta(\bar{\nu}_1^a - \bar{\nu}_2^a) - \bar{\nu}_2^a + \frac{1}{2}(\eta + 1)\eta \left(\frac{\mu_1^a - \mu_2^a}{\sigma^a} \right)^2 - r = 0$$

in $(0, \infty)$, and let $\beta_{a,b}$ be the unique solution of Eq. (2.8). Using similar arguments to those used to prove Lemma 4.4, one can show that $\beta_{a,b} < \beta_a$. Let $p_{a,c}^* = \frac{\beta_a(s - g_2^a - g_2^c)}{(\beta_a + 1)(g_1^a + g_1^c - s) + \beta_a(s - g_2^a - g_2^c)}$ be the optimal cut-off of DM2. By Theorem 2.6, the optimal expected payoff of DM2 is given by

$$\begin{aligned} U_{DM2}(p_0) &= V_{p_{a,c}^*}^{a,c}(p_0, p_0) \\ &= \begin{cases} s & \text{if } p_0 \leq p_{a,c}^*, \\ p_0(g_1^a + g_1^c) + (1 - p_0)(g_2^a + g_2^c) + C_{\beta_a}(1 - p_0)\left(\frac{1 - p_0}{p_0}\right)^{\beta_a} & \text{if } p_0 > p_{a,c}^*, \end{cases} \end{aligned}$$

$$\text{where } C_{\beta_a} = \frac{s - g_2^a - g_2^c - p_{a,c}^*(g_1^a - g_1^c - g_2^a - g_2^c)}{(1 - p_{a,c}^*) \left(\frac{1 - p_{a,c}^*}{p_{a,c}^*} \right)^{\beta_a}}.$$

By Theorem 2.6, the expected payoff of DM1, using the same cut-off point $p_{a,c}^*$ is given by

$$V_{p_{a,c}^*}^{a,b,c}(p_0, p_0) = \begin{cases} s & \text{if } p_0 \leq p_{a,c}^*, \\ p_0(g_1^a + g_1^c) + (1 - p_0)(g_2^a + g_2^c) + C_{\beta_{a,b}}(1 - p_0)\left(\frac{1 - p_0}{p_0}\right)^{\beta_{a,b}} & \text{if } p_0 > p_{a,c}^*, \end{cases}$$

$$\text{where } C_{\beta_{a,b}} = \frac{s - g_2^a - g_2^c - p_{a,c}^*(g_1^a - g_1^c - g_2^a - g_2^c)}{(1 - p_{a,c}^*) \left(\frac{1 - p_{a,c}^*}{p_{a,c}^*} \right)^{\beta_{a,b}}}.$$

Since $\beta_{a,b} < \beta_a$, it follows that if $p_0 > p_{a,c}^*$, then $U_{DM2}(p_0) = V_{p_{a,c}^*}^{a,c}(p_0, p_0) < V_{p_{a,c}^*}^{a,b,c}(p_0, p_0) \leq U_{DM1}(p_0)$, as desired.

□

4.6. Optimism vs Pessimism.

Proof of Theorem 2.8. Substituting $q_0 = p_0 + \epsilon$ in Eq. (2.7) yields

$$V_{p_0}(\epsilon) = p_0 g_1 + (1 - p_0) g_2 + p_0 (s - g_1) \left(\frac{p^*}{1 - p^*} \right)^{\alpha+1} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha+1} \\ + (1 - p_0) (s - g_2) \left(\frac{p^*}{1 - p^*} \right)^{\alpha} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha},$$

where α is the unique solution of (2.2) in $(0, \infty)$.

Simple algebraic manipulations yield

$$\begin{aligned} V_{p_0}(\epsilon) &= p_0 g_1 + (1 - p_0) g_2 + p_0 (s - g_1) \left(\frac{p^*}{1 - p^*} \right)^{\alpha+1} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha+1} \\ &\quad + (1 - p_0) (s - g_2) \left(\frac{p^*}{1 - p^*} \right)^{\alpha} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha} \\ &= [p_0 g_1 + (1 - p_0) g_2] + \left(\frac{p^*}{1 - p^*} \right)^{\alpha} \cdot \\ &\quad \cdot \left[p_0 (s - g_1) \frac{p^*}{1 - p^*} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha+1} + (1 - p_0) (s - g_2) \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha} \right] \\ &= [p_0 g_1 + (1 - p_0) g_2] + \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha} \left(\frac{s - g_2}{g_1 - s} \right)^{\alpha} \cdot \\ &\quad \cdot \left[-p_0 (s - g_2) \frac{\alpha}{\alpha + 1} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha+1} + (1 - p_0) (s - g_2) \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha} \right] \\ &= [p_0 g_1 + (1 - p_0) g_2] + \left(\frac{\alpha}{\alpha + 1} \right)^{\alpha} \left(\frac{s - g_2}{g_1 - s} \right)^{\alpha} \frac{1}{p_0 (s - g_2)} \cdot \\ &\quad \cdot \left[-\frac{\alpha}{\alpha + 1} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha+1} + \frac{1 - p_0}{p_0} \left(\frac{1 - p_0 - \epsilon}{p_0 + \epsilon} \right)^{\alpha} \right]. \end{aligned}$$

For every $x \geq 0$ define $W(x) = V_p(x) - V_p(-x)$. This is the difference between the payoff of an optimist and the payoff of a pessimist. Straightforward calculations show that

$$W'(x) = \frac{\alpha x}{p} \left[- \left(\frac{1 - (p + x)}{p + x} \right)^{\alpha} \frac{1}{(p + x)^2 (1 - (p + x))} + \left(\frac{1 - (p - x)}{p - x} \right)^{\alpha} \frac{1}{(p - x)^2 (1 - (p - x))} \right],$$

so that $W(0) = W'(0) = 0$. Suppose $\alpha > 1$. Since $\frac{1}{(p+x)^3} < \frac{1}{(p-x)^3}$, and $\left(\frac{1-(p+x)}{p+x} \right)^{\alpha-1} < \left(\frac{1-(p-x)}{p-x} \right)^{\alpha-1}$, we get $W'(x) > 0$ for every $x > 0$ such that $p^* < p \pm x \leq 1$, and so $W(x) > 0$; an optimist will fare better.

If $0 < \alpha < 1$, it is easy to verify that $p^* < \frac{\alpha+2}{3}$. Since the function $\left(\frac{1-y}{y} \right)^{\alpha} \frac{1}{y^2(1-y)}$ decreases for $0 < y < \frac{\alpha+2}{3}$ and increases for $\frac{\alpha+2}{3} < y < 1$, it follows that for every $x > 0$ such that $p^* < p \pm x \leq \frac{\alpha+2}{3}$ we get $W'(x) > 0$; an optimist will fare better, and for every $x > 0$ such that $\frac{\alpha+2}{3} < p \pm x$ we get $W'(x) < 0$, and so $W(x) < 0$; a pessimist will fare better. \square

5. FIGURES

In Figure 1 we depict a generic path of the Lévy payoff process ($Y(t)$). The contribution of the compound Poisson process (the jumps) appears in Figure 2. The posterior belief given the observations appears in Figure 3. Finally, Figure 4 shows the time-dependent cut-off, as well as the continuous part of the process (the Brownian motion with drift).

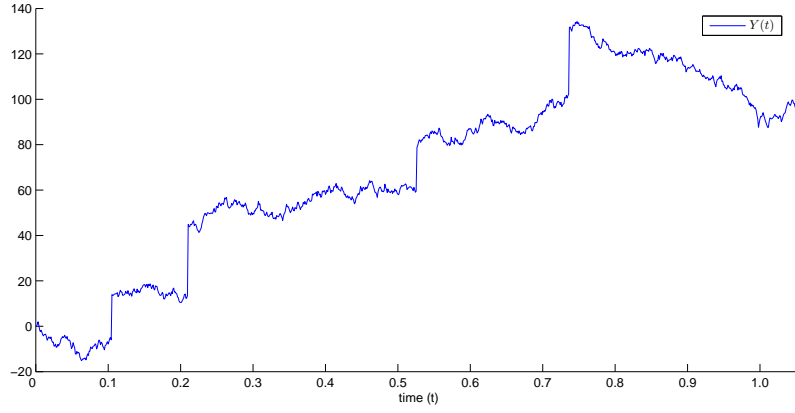


FIGURE 1. The payoff process

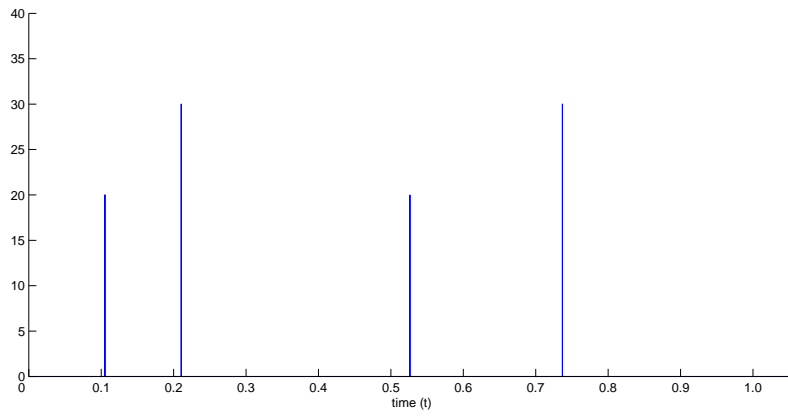


FIGURE 2. The Poisson arrivals

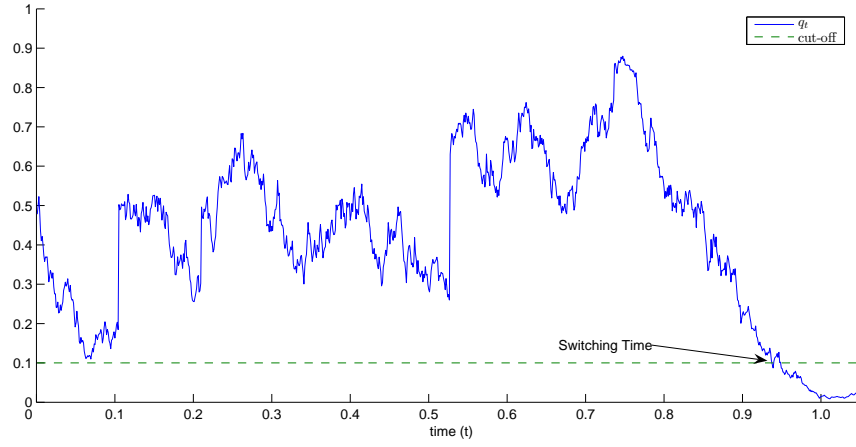


FIGURE 3. The posterior process

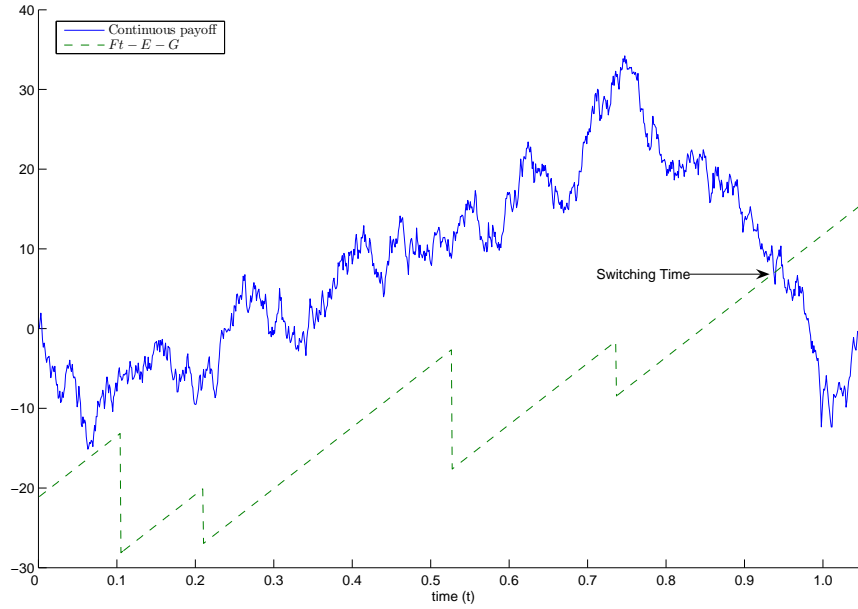


FIGURE 4. The strategy description

REFERENCES

1. Applebaum D.(2004) Lévy Processes and Stochastic Calculus. Cambridge.
2. Babcock L. and Loewenstein G. (1997) Explaining Bargaining Impasse: The Role of Self-Serving Biases. *Journal of Economic Perspectives*, **11**, 109-126.
3. Baumhart R. (1968) *An Honest Profit*. Prentice Hall.

4. Bergemann D. and Välimäki J. (2006). Bandit Problems. HECER discussion paper number 93.
5. Berry D. A. and Fristedt, B. (1985). Bandit Problems: Sequential Allocation of Experiments. Chapman and Hall, New York.
6. Bertoin J. (1996) Lévy Processes. Cambridge University Press. Cambridge.
7. Besanko D. and Wu J. (2008) The Impact of Market Structure on R&D Competition with Learning. Preprint.
8. Bolton P. and Harris C. (1999) Strategic Experimentation. *Econometrica*, **67**, 349-374.
9. Borodin A.N. and Salminen P. (1996) Handbook of Brownian motion - facts and formulae. Basel, Birkhauser.
10. Camerer C.F. and Lovo D. (1999) Overconfidence and Excess Entry: An Experimental Approach. *The American Economic Review*, **89**, 306-318.
11. Carr P. and Wu L. (2004) Time-Changed Lévy Processes and Option Pricing. *Journal of Financial Economics*, **71**, 113-141.
12. Cohen A. and Solan E. (2004) Two armed Lévy bandit with infinite Lévy measure, preprint.
13. Cross P. (1977) Not can, But will College Teaching be Improved? *New Directions for Higher Education*, **17**, 1-15.
14. Friedman A. (1969) Partial Differential Equations, Holt, Rinehart & Winston.
15. Geman H., Madan D. and Yor M. (2001) Time Changes for Lévy Processes. *Mathematical Finance*, **17**, 79-96.
16. Gittins J.C. and Jones D.M (1979) A Dynamic Allocation Index for the Discounted Multi-armed Bandit Problem. *Biometrika*, **66**, 561-565.
17. Heifetz A., Shannon C. and Spiegel Y. (2007) The Dynamic Evolution of Preferences. *Economic Theory*, **32**, 251-286.
18. Karatzas I. (1984) Gittins indices in the dynamic allocation problem for diffusion processes. *Annals of Probability*, **12**, 173-192.
19. Kaspi H. and Mandelbaum A. (1995) Lévy Bandits: Multi-Armed Bandits Driven by Lévy Processes. *Ann. Appl. Probab.*, **5**, 541-565.
20. Keller G., Rady S. and Cripps M. (2005) Strategic Experimentation with Exponential Bandits. *Econometrica*, **73**, 39-68.
21. Klein N. and Rady S. (2008) Negatively Correlated Bandits, under revision.
22. Larwood L. and Whittaker W. (1977) Managerial Myopia: Self-Serving Biases in Organizational Planning. *Journal of Applied Psychology*, **62**, 194-198.
23. Moscarini G. and Squintani F. (2004) Competitive Experimentation with Private Information. *Cowles Foundation DP 1489*.
24. Roberts K. and Weitzman M.L. (1981) Funding Criteria for Research, Development, and Exploration Projects. *Econometrica*, **49**, 1261-1288.
25. Rothschild M. (1974) A Two-Armed Bandit Theory of Market Pricing. *Journal of Economic Theory*, **9**, 185-202.
26. Svenson O. (1981) Are We all Less Risky and More Skillful than our Fellow Drivers? *Acta Psychologica*, **47**, 143-148.
27. Weinstein N.D. (1980) Unrealistic Optimism about Future Life Events. *Journal of Personality and Social Psychology*, **39**, 806-820.
28. Weitzman M.L. (1979) Optimal Search for the Best Alternative. *Econometrica*, **47**, 641-654.